

# HOMOGENIZATION OF PARABOLIC EQUATIONS WITH A CONTINUUM OF SPACE AND TIME SCALES\*

HOUMAN OWHADI<sup>†</sup> AND LEI ZHANG<sup>‡</sup>

**Abstract.** This paper addresses the issue of the homogenization of linear divergence form parabolic operators in situations where no ergodicity and no scale separation in time or space are available. Namely, we consider divergence form linear parabolic operators in  $\Omega \subset \mathbb{R}^n$  with  $L^\infty(\Omega \times (0, T))$ -coefficients. It appears that the inverse operator maps the unit ball of  $L^2(\Omega \times (0, T))$  into a space of functions which at small (time and space) scales are close in  $H^1$  norm to a functional space of dimension  $n$ . It follows that once one has solved these equations at least  $n$  times it is possible to homogenize them both in space and in time, reducing the number of operation counts necessary to obtain further solutions. In practice we show under a Cordes-type condition that the first order time derivatives and second order space derivatives of the solution of these operators with respect to caloric coordinates are in  $L^2$  (instead of  $H^{-1}$  with Euclidean coordinates). If the medium is time-independent, then it is sufficient to solve  $n$  times the associated elliptic equation in order to homogenize the parabolic equation.

**Key words.** multiscale problem, compensation, homogenization, upscaling, compression

**AMS subject classifications.** Primary, 34E13, 35B27; Secondary, 35B26, 68P30, 60F05, 35B05

**DOI.** 10.1137/060670420

**1. Introduction and main results.** Let  $\Omega$  be a bounded and convex domain of class  $C^2$  of  $\mathbb{R}^n$ . Let  $T > 0$ . Consider the following parabolic PDE:

$$(1.1) \quad \begin{cases} \partial_t u = \operatorname{div} (a(x, t) \nabla u(x, t)) + g & \text{in } \Omega \times (0, T), \\ u(x, t) = 0 & \text{for } (x, t) \in (\partial\Omega \times (0, T)) \cup (\Omega \times \{t = 0\}). \end{cases}$$

Write  $\Omega_T := \Omega \times (0, T)$ .  $g$  is a function in  $L^2(\Omega_T)$ .  $(x, t) \rightarrow a(x, t)$  is a mapping from  $\Omega_T$  into the space of symmetric positive definite matrices with entries in  $L^\infty(\Omega_T)$ . Assume  $a$  to be uniformly elliptic on the closure of  $\Omega_T$ . This paper addresses the issue of the homogenization of (1.1) in space and time in situations where scale separation and ergodicity at small scales are not available (see [16], [43], [4] for an introduction to classical homogenization theory).

We will introduce in subsection 1.1 theorems establishing under Cordes-type conditions the increase of regularity of solutions of (1.1) when derivatives are taken with respect to caloric coordinates instead of Euclidean coordinates. In subsections 1.2 and 1.3 these results will be used to homogenize (1.1) in space and in time. More precisely, assume  $a$  to be written on a fine tessellation with  $N$  degrees of freedom. If  $a$  is time-independent, then by solving  $n$  times an elliptic boundary value problem associated to (1.1) (at a cost of  $O(N(\ln N)^{n+3})$  operations, using the hierarchical matrix method [11] or in  $O(N)$  operations using iterative methods (see [76], [77]) it

---

\*Received by the editors September 22, 2006; accepted for publication (in revised form) August 10, 2007; published electronically December 12, 2007.

<http://www.siam.org/journals/sinum/46-1/67042.html>

<sup>†</sup>Department of Applied and Computational Mathematics, Control, and Dynamical Systems, California Institute of Technology MC 217-50, Pasadena, CA 91125 (owhadi@caltech.edu).

<sup>‡</sup>Department of Applied and Computational Mathematics, California Institute of Technology, MC 217-50, Pasadena, CA 91125 (zhanglei@acm.caltech.edu).

is possible to approximate the solutions of (1.1) by solving an homogenized operator with  $N^\alpha$  degrees of freedom ( $\alpha < 1$ ,  $\alpha = 0.2$ , for instance) or with a fixed number  $M$  of degrees of freedom (numerical experiments given at the end of this paper have been conducted with  $N = 16641$  and  $M = 9$ ). Therefore, in the case of time-independent  $a$ , it is possible to reduce the number of operations even if we solve the problem (1.1) only once by using large time steps. If  $a$  is also characterized by a continuum of time scales, then the method presented here does not reduce the number of operation counts necessary to solve (1.1) only one time. However, if one needs to solve (1.1)  $K$  ( $K > n$ ) times (with different right-hand sides), then by solving (1.1)  $n$  times it is possible to obtain an approximation of the solutions of (1.1) by solving an homogenized (in space and time) parabolic equation written on a coarse tessellation with coarse time steps. From a broader point of view, this paper suggests that, when one is trying to solve the linear problem  $Ax = y$  with many different  $y$  (where  $A$  is a matrix associated to a divergence form operator), one should never compute  $A^{-1}$ .

In [60] the authors have addressed the issue of the homogenization of divergence form elliptic equations with a continuum of space scales; however, this paper is not a technical extension of the previous one. Indeed the presence of temporal scales adds strong difficulties through the fact that for parabolic equations finite element methods are no longer optimal. The main difference when the coefficients are rough in both space and time lies in the fact that we have to introduce time-dependent elements and use dual methods for controlling the numerical homogenization errors instead of simply relying on the increased regularity. Furthermore the numerical scheme has to be implicit in time.

Numerical results given in section 3 show that the method can be implemented and work in practice. The idea to use a change of coordinates for a finite element method for divergence form elliptic equations with rough coefficients can be traced back to the work of Babuška et al. [9, 8] (in dimension two with coefficients depending only on one parameter). The idea to use a global change of coordinates to homogenize transport equations is currently being implemented in the industry and has been shown to give accurate results when local methods fail. For instance, we refer to [75], [74] for reservoir modeling in geophysics; in this situation the porosity of the medium is time-independent, and one has to solve an elliptic equation only at  $t = 0$  to upscale the transport equations. Observe also that using an implicit method to solve divergence form parabolic equations with time-independent rough coefficients implies solving an elliptic problem at each time step. The method presented in this paper implies that these elliptic problems can be solved with an “homogenized” operator leading to a sharp improvement of the computational cost.

Although we use the term “homogenization” in this paper for lack of a better word, the method presented in this paper is more general than homogenization. Homogenization is typically associated with the limit points of a sequence of operators. Here we are speaking of a fixed operator.

**1.1. Compensation phenomenon.** Let  $F$  be the solution of the following parabolic equation:

$$(1.2) \quad \begin{cases} \partial_t F = \operatorname{div}(a(x, t) \nabla F(x, t)) & \text{in } \Omega_T, \\ F(x, t) = x & \text{for } (x, t) \in (\partial\Omega \times (0, T)), \\ \operatorname{div}(a(x, 0) \nabla F(x, 0)) = 0 & \text{in } \Omega. \end{cases}$$

By (1.2) we mean that  $F := (F_1, \dots, F_n)$  is a  $n$ -dimensional vector field such that each of its entries satisfies

$$(1.3) \quad \begin{cases} \partial_t F_i = \operatorname{div} (a(x, t) \nabla F_i(x, t)) & \text{in } \Omega_T, \\ F_i(x, t) = x_i & \text{for } (x, t) \in (\partial\Omega \times (0, T)), \\ \operatorname{div} (a(x, 0) \nabla F_i(x, 0)) = 0 & \text{in } \Omega. \end{cases}$$

In the literature  $F$  is also called a caloric coordinate associated to (1.1). Observe that if  $a$  is time-independent, then  $F$  is the solution of an elliptic boundary value problem.

DEFINITION 1.1. *Write*

$$(1.4) \quad \sigma := {}^t \nabla F a \nabla F.$$

Write  $\beta_\sigma$  as the Cordes parameter associated to  $\sigma$  defined by

$$(1.5) \quad \beta_\sigma := \operatorname{esssup}_{(x,t) \in \Omega_T} \left( n - \frac{(\operatorname{Trace}[\sigma])^2}{\operatorname{Trace}[{}^t \sigma \sigma]} \right).$$

Observe that since

$$(1.6) \quad \beta_\sigma = \operatorname{esssup}_{(x,t) \in \Omega_T} \left( n - \frac{(\sum_{i=1}^n \lambda_{i,\sigma(x,t)})^2}{\sum_{i=1}^n \lambda_{i,\sigma(x,t)}^2} \right),$$

where  $(\lambda_{i,M})$  denotes the eigenvalues of  $M$ ,  $\beta_\sigma$  is a measure of the anisotropy of  $\sigma$ . The greater  $\beta_\sigma$  is, the more anisotropic  $\sigma$  is, and  $\beta_\sigma = 0$  corresponds to an isotropic  $\sigma$ .

**1.1.1. Time-independent medium.** In this subsection we assume that  $a$  does not depend on time  $t$ . Write for  $p \geq 2$ ,  $W_D^{2,p}$  ( $D$  for Dirichlet boundary condition) the Banach space  $W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega)$ . Equip  $W_D^{2,p}(\Omega)$  with the norm

$$(1.7) \quad \|v\|_{W_D^{2,p}(\Omega)}^p := \int_\Omega \left( \sum_{i,j} (\partial_i \partial_j v)^2 \right)^{\frac{p}{2}}.$$

Equip the space  $L^p(0, T, W_D^{2,p}(\Omega))$  with the norm

$$(1.8) \quad \|v\|_{L^p(0,T,W_D^{2,p}(\Omega))}^p = \int_0^T \int_\Omega \left( \sum_{i,j} (\partial_i \partial_j v)^2 \right)^{\frac{p}{2}} dx dt.$$

THEOREM 1.1. *Assume that  $\partial_t a \equiv 0$ ,  $g \in L^2(\Omega_T)$ ,  $\Omega$  is convex,  $\beta_\sigma < 1$ , and  $(\operatorname{Trace}[\sigma])^{\frac{n}{4}-1} \in L^\infty(\Omega)$ ; then  $u \circ F^{-1} \in L^2(0, T, W_D^{2,2}(\Omega))$  and*

$$(1.9) \quad \|u \circ F^{-1}\|_{L^2(0,T,W_D^{2,2}(\Omega))} \leq \frac{C}{1 - \beta_\sigma^{\frac{1}{2}}} \|g\|_{L^2(\Omega_T)}.$$

Remark 1.1. The constant  $C$  can be written

$$C = \frac{C_n}{(\lambda_{\min}(a))^{\frac{n}{4}}} \|(\operatorname{Trace}[\sigma])^{\frac{n}{4}-1}\|_{L^\infty(\Omega)}.$$

Throughout this paper, we write

$$(1.10) \quad \lambda_{\min}(a) := \inf_{(x,t) \in \Omega_T} \inf_{l \in \mathbb{R}^n, |l|=1} {}^t l a(x, t) l.$$

All of the results are stated in a finite time interval  $[0, T]$ .  $T$  is just any given finite time, and it is easy to observe that if  $\|g\|_{L^2(\Omega_\infty)} < \infty$ , then one can take  $T = \infty$ .

*Remark 1.2.* According to Theorem 1.1 although the second order derivatives of  $u$  with respect to Euclidean coordinates are only in  $L^2(0, T, H^{-1}(\Omega))$ , they are in  $L^2(\Omega_T)$  with respect to harmonic coordinates. The proof of that theorem (given in subsection 2.1.1) lies on the fact  $u \circ F^{-1}$  satisfies nondivergence form parabolic equations, and the solutions of these equations are known to have a higher regularity under a Cordes-type condition on the coefficients and under a convexity assumption on the domain. One can obtain a similar result without the convexity assumption on the domain (Theorem 2.6), but the Cordes-type condition on the coefficients would have to be stronger.

*Remark 1.3.* Observe that if  $a$  is time-independent, then  $F$  and  $\sigma$  are time-independent, and  $F$  is the solution of the following elliptic problem:

$$(1.11) \quad \begin{cases} \operatorname{div} a \nabla F = 0 & \text{in } \Omega, \\ F(x) = x & \text{on } \partial\Omega. \end{cases}$$

In dimension one  $F$  is trivially a homeomorphism. In dimension two this property follows from the convexity of the domain [3] (even with  $a_{i,j} \in L^\infty(\Omega)$ ), [6] (one can also deduce from [6] that for  $n = 2$ , if  $a$  is smooth, then the conditions  $\beta_\sigma < 1$  and  $(\operatorname{Trace}[\sigma])^{-1} \in L^\infty(\Omega)$  are satisfied). In dimension three and higher  $F$  can be nonbijective even if  $a$  is smooth; we refer to [6], [22]. However in dimension three the assumption  $(\operatorname{Trace}[\sigma])^{\frac{n}{4}-1} \in L^\infty(\Omega_T)$  implies that  $F$  is a homeomorphism. If  $n \geq 4$ , we need to assume that  $F$  is a homeomorphism to prove the theorem.

*Remark 1.4.* In fact the condition  $(\operatorname{Trace}(\sigma))^{-1} \in L^p(\Omega_T)$ , for  $p < \infty$  depending on  $n$ , is sufficient to obtain Theorem 1.1 and the following compensation theorems. For the sake of clarity this paper has been restricted to  $(\operatorname{Trace}(\sigma))^{-1} \in L^\infty(\Omega_T)$ .

*Remark 1.5.* Write

$$(1.12) \quad \mu_\sigma := \operatorname{esssup}_{\Omega_T} \frac{\lambda_{\max}(\sigma)}{\lambda_{\min}(\sigma)}.$$

It is easy to check that  $\mu_\sigma$  is bounded by an increasing function of  $(1 - \beta_\sigma)^{-1}$ , and in dimension two  $\beta_\sigma < 1$  is equivalent to  $\mu_\sigma < \infty$ .

*Remark 1.6.* Theorem 1.1 has been called the compensation phenomenon because the composition by  $F^{-1}$  increases the regularity of  $u \in L^2(0, T, H_0^1(\Omega))$ . The choice of this name has been motivated by Murat and Tartar's work on  $H$ -convergence [54] which is also based on a regularization property called compensated compactness or the div-curl lemma introduced in the 1970s by Murat and Tartar [53], [70] (we also refer to [26] for refinements of the div-curl lemma).

The compensation phenomena presented in this subsection can be observed numerically. We consider the time-independent site percolation, and the value of  $a$  is set to be equal to 1 or 100 with probability 1/2 on each triangle of a fine mesh characterized by 16641 nodes and 32768 triangles. (1.1) has been solved numerically on that mesh with  $g = 1$ .  $u$ ,  $u \circ F^{-1}$ ,  $\partial_x u$ , and  $\partial_x(u \circ F^{-1})$  are plotted at time  $t = 1$  in Figure 1.1.

It is easy to check that if  $a = e(x)S(t)$ , where  $e$ , is a time-independent symmetric uniformly elliptic matrix with  $L^\infty(\Omega)$  entries and  $S$  is a smooth uniformly positive function, then the results given in this subsection and the homogenization schemes of subsection 1.2 remain valid with the time-independent harmonic coordinates associated to  $e$ , i.e., solution of  $-\operatorname{div} e \nabla F = 0$ .

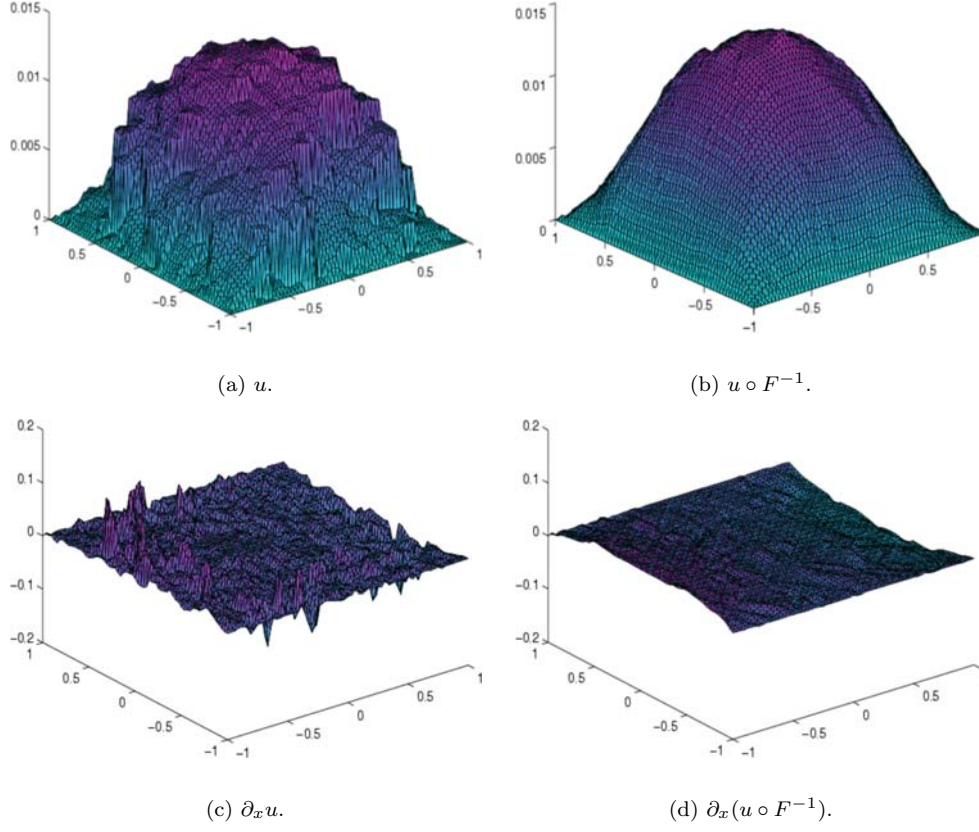


FIG. 1.1.  $u$ ,  $u \circ F^{-1}$ ,  $\partial_x u$ , and  $\partial_x(u \circ F^{-1})$  at time  $t = 1$  for the time-independent site percolating medium.

**1.1.2. Medium with a continuum of time scales.** In this subsection the entries of  $a$  are merely in  $L^\infty(\Omega_T)$ . We need to introduce the following Cordes-type condition.

CONDITION 1.1. *We say that Condition 1.1 is satisfied if and only if there exists  $\delta \in (0, \infty)$  and  $\epsilon > 0$  such that*

$$(1.13) \quad \operatorname{esssup}_{\Omega_T} \frac{\delta^2 \operatorname{Trace}[{}^t \sigma \sigma] + 1}{(\delta \operatorname{Trace}[\sigma] + 1)^2} \leq \frac{1}{n + \epsilon}.$$

*Remark 1.7.* This remark is due to Gyrya and Berlyand. Equation (1.13) can be written in terms of the Cordes parameter by

$$(1.14) \quad \operatorname{esssup}_{\Omega_T} \left\{ n - \frac{(\operatorname{Trace}(\sigma) + \delta^{-1})^2}{\operatorname{Trace}(\sigma^t \sigma) + \delta^{-2}} \right\} \leq -\epsilon.$$

An alternative way of interpreting the Cordes-type condition is by rewriting (1.14) as follows:

$$(1.15) \quad \beta_{\hat{\sigma}(\delta)} = \operatorname{esssup}_{\Omega_T} \left\{ (n + 1) - \frac{(\operatorname{Trace}(\sigma) + \delta^{-1})^2}{\operatorname{Trace}(\sigma^t \sigma) + \delta^{-2}} \right\} \leq 1 - \epsilon,$$

where

$$\hat{\sigma}(\delta) = \begin{bmatrix} \delta^{-1} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \sigma & \\ 0 & & & \end{bmatrix}.$$

Write

$$(1.16) \quad z_\sigma := \text{esssup}_{\Omega_T} n \frac{\text{Trace}[\sigma]}{(\text{Trace}[\sigma])^2}.$$

Observe that  $z_\sigma$  is a measure of anisotropy of  $\sigma$ , in particular,  $1 \leq z_\sigma \leq n$  and  $z_\sigma = 1$  if  $\sigma$  is isotropic. Write

$$(1.17) \quad y_\sigma := \|\text{Trace}[\sigma]\|_{L^\infty(\Omega_T)} \|(\text{Trace}[\sigma])^{-1}\|_{L^\infty(\Omega_T)}.$$

PROPOSITION 1.1. *If  $\|\text{Trace}[\sigma]\|_{L^\infty(\Omega_T)} < \infty$  and  $\|(\text{Trace}[\sigma])^{-1}\|_{L^\infty(\Omega_T)} < \infty$ , then Condition 1.1 is satisfied with*

$$(1.18) \quad \delta := n \|(\text{Trace}[\sigma])^{-1}\|_{L^\infty(\Omega_T)}$$

and with  $\epsilon := \frac{2ny_\sigma - n}{2ny_\sigma^2}$  provided that  $z_\sigma \leq 1 + \frac{\epsilon}{n}$ .

Remark 1.8. Observe that in dimension one  $z_\sigma = 1$ , and thus for  $n = 1$  Condition 1.1 is satisfied if  $\text{Trace}[\sigma] \in L^\infty(\Omega_T)$  and  $(\text{Trace}[\sigma])^{-1} \in L^\infty(\Omega_T)$ .

We have the following compensation theorem for a time-dependent medium.

THEOREM 1.2. *Assume that  $\Omega$  is convex, and Condition 1.1 is satisfied, then  $u \circ F^{-1} \in L^2(0, T, W_D^{2,2}(\Omega))$ ,  $\partial_t(u \circ F^{-1}) \in L^2(\Omega_T)$ , and*

$$(1.19) \quad \|u \circ F^{-1}\|_{L^2(0, T, W_D^{2,2}(\Omega))} + \|\partial_t(u \circ F^{-1})\|_{L^2(\Omega_T)} \leq C \|g\|_{L^2(\Omega_T)},$$

where  $C$  depends on  $\Omega$ ,  $n$ ,  $\delta$ , and  $\epsilon$ .

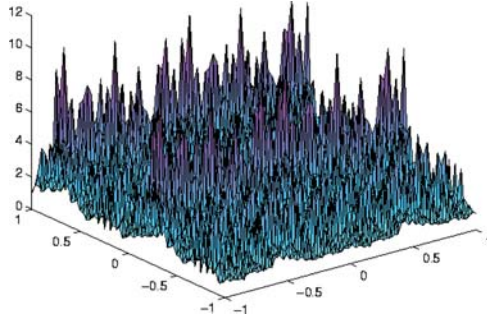
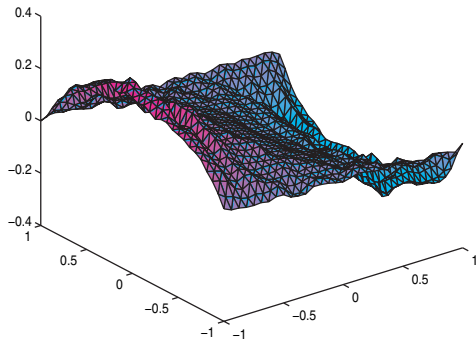
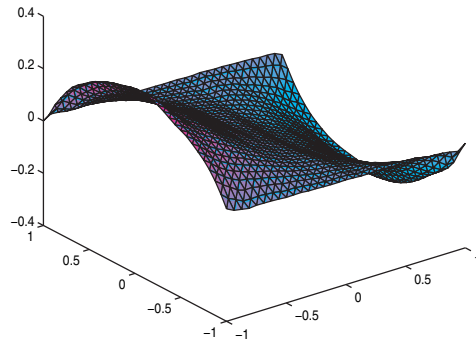
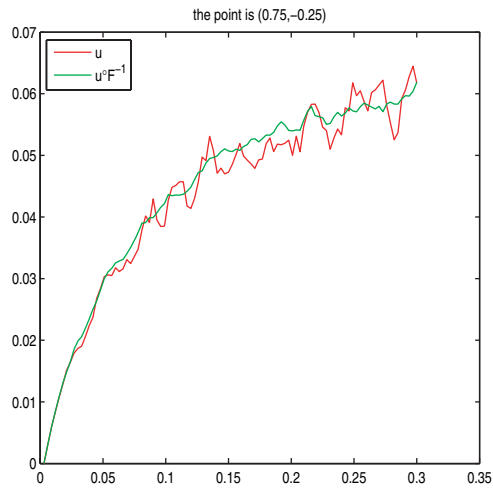
Remark 1.9. According to Theorem 1.2 although the second order space derivatives and first order time derivatives of  $u$  with respect to Euclidean coordinates are only in  $L^2(0, T, H^{-1}(\Omega))$ , they are in  $L^2(\Omega_T)$  with respect to caloric coordinates. As in the time-independent, the proof (given in subsection 2.1.2) lies on the fact that  $u \circ F^{-1}$  satisfies a nondivergence form parabolic equation.

The compensation phenomena can be observed numerically. We consider in dimension  $n = 2$

$$(1.20) \quad a(x, y, t) = \frac{1}{6} \left( \sum_{i=1}^5 \frac{1.1 + \sin(2\pi x'/\epsilon_i)}{1.1 + \sin(2\pi y'/\epsilon_i)} + \sin(4x'^2 y'^2) + 1 \right),$$

with  $x' = x + \sqrt{2}t$ ,  $y' = y - \sqrt{2}t$ ,  $\epsilon_1 = \frac{1}{5}$ ,  $\epsilon_2 = \frac{1}{13}$ ,  $\epsilon_3 = \frac{1}{17}$ ,  $\epsilon_4 = \frac{1}{31}$ , and  $\epsilon_5 = \frac{1}{65}$ . This medium has been plotted in Figure 1.2 at time 0 (observe that  $\lambda_{\max}(a)/\lambda_{\min}(a) \sim 100$ ).

Equation (1.1) has been solved numerically on that mesh with  $g \equiv 1$  on the fine mesh with 16641 nodes and 32768 triangles. Figure 1.3 shows  $\partial_x u$  and  $\partial_x(u \circ F^{-1})$  at time 0.3. In Figures 1.4 and 1.5, the value of  $x_0$  is set to  $x_0 := (0.75, -0.25)$ , and the curves  $t \rightarrow u(x_0, t)$ ,  $u \circ F^{-1}(x_0, t)$ ,  $\nabla u(x_0, t)$ , and  $\nabla u \circ F^{-1}(x_0, t)$  are plotted from  $t = 0$  to  $t = 0.3$ .

FIG. 1.2.  $a$  at time 0.(a)  $\partial_x u$ .(b)  $\partial_x(u \circ F^{-1})$ .FIG. 1.3.  $\partial_x u$  and  $\partial_x(u \circ F^{-1})$  at time  $t = 0.3$  for the multiscale trigonometric time-dependent medium.(a)  $u$  and  $u \circ F^{-1}$ .FIG. 1.4.  $t \rightarrow u(x_0, t), u \circ F^{-1}(x_0, t)$  from  $t = 0$  to  $t = 0.3$ , with  $x_0 := (0.75, -0.25)$ .

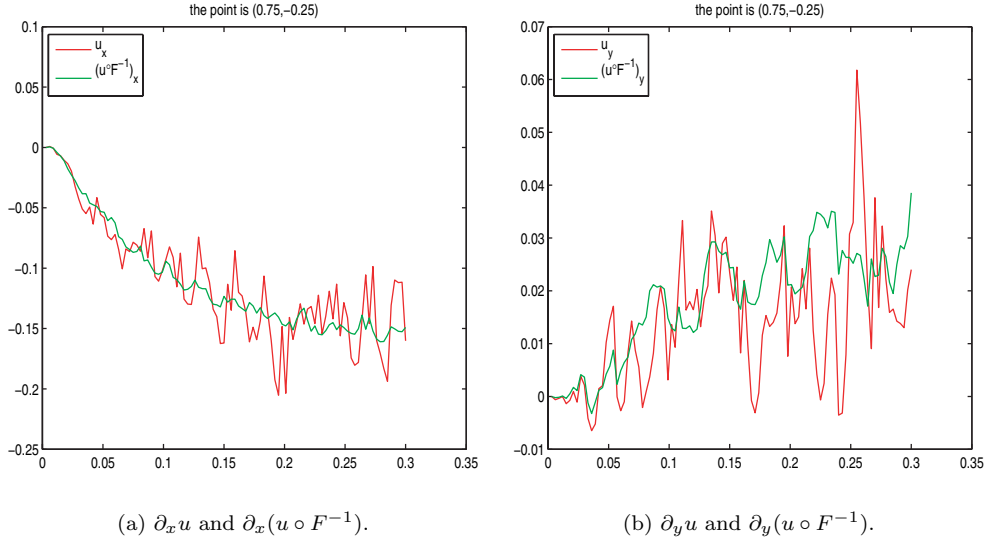


FIG. 1.5.  $t \rightarrow \nabla u(x_0, t), \nabla u \circ F^{-1}(x_0, t)$  from  $t = 0$  to  $t = 0.3$ , with  $x_0 := (0.75, -0.25)$ .

**1.2. Numerical homogenization in space.** Let  $X^h$  be a finite-dimensional subspace of  $H_0^1(\Omega) \cap W^{1,\infty}(\Omega)$  ( $W^{1,\infty}$  is the usual space of uniformly Lipschitz continuous functions) with the following approximation property: There exists a constant  $C_X$  such that for all  $f \in W_D^{2,2}(\Omega)$

$$(1.21) \quad \inf_{v \in X^h} \|f - v\|_{H_0^1(\Omega)} \leq C_X h \|f\|_{W_D^{2,2}(\Omega)}.$$

It is known and easy to check that the set of piecewise linear functions on a triangulation of  $\Omega$  satisfies condition (1.21) provided that the length of the edges of the triangles are bounded by  $h$  ( $C_X$  in (1.21) being given by the aspect ratio of the triangles).

For media characterized by a continuum of time scales we will consider twice differentiable elements satisfying the following usual inverse inequalities (see section 1.7 of [30]): For  $v \in X^h$ ,

$$(1.22) \quad \|v\|_{W_D^{2,2}(\Omega)} \leq C_X h^{-1} \|v\|_{H_0^1(\Omega)}.$$

and

$$(1.23) \quad \|v\|_{H_0^1(\Omega)} \leq C_X h^{-1} \|v\|_{L^2(\Omega)}.$$

In this paper we will use splines to ensure that condition (1.22) is satisfied (observe that it requires the quasi uniformity of the (coarse) mesh, i.e., a bound on the aspect ratio of the (coarse) triangles).

For  $t \in (0, T)$  let us define

$$(1.24) \quad V_h(t) := \{\varphi \circ F(x, t) : \varphi \in X^h\}.$$

Write  $L^2(0, T; H_0^1(\Omega))$ , the usual Sobolev space associated to the norm

$$(1.25) \quad \|v\|_{L^2(0, T; H_0^1(\Omega))}^2 := \int_0^T \|v(\cdot, t)\|_{H_0^1(\Omega)}^2 dt.$$



Define  $Y_T^h$ , the subspace of  $L^2(0, T; H_0^1(\Omega))$ , as

$$(1.26) \quad Y_T^h := \{v \in L^2(0, T; H_0^1(\Omega)) : v(x, t) \in V_h(t) \forall t \in [0, T]\}.$$

Write  $u_h$ , the solution in  $Y_T^h$  of the following system of ordinary differential equations:

$$(1.27) \quad \begin{cases} (\psi, \partial_t u_h)_{L^2(\Omega)} + a[\psi, u_h] = (\psi, g)_{L^2(\Omega)} & \text{for all } t \in (0, T) \text{ and } \psi \in V_h(t), \\ u_h(x, 0) = 0. \end{cases}$$

Write

$$(1.28) \quad a[v, w] := \int_{\Omega} {}^t \nabla v(x, t) a(x, t) \nabla w(x, t) dx.$$

**1.2.1. Time-independent domain.** We have the following theorem.

**THEOREM 1.3.** *Assume that  $\partial_t a \equiv 0$ ,  $\Omega$  is convex,  $\beta_{\sigma} < 1$ , and  $(\text{Trace}[\sigma])^{-1} \in L^{\infty}(\Omega_T)$ ; then*

$$(1.29) \quad \| (u - u_h)(\cdot, T) \|_{L^2(\Omega)} + \| u - u_h \|_{L^2(0, T; H_0^1(\Omega))} \leq Ch \| g \|_{L^2(\Omega_T)}.$$

*Remark 1.10.* The constant  $C$  depends on  $C_X$ ,  $n$ ,  $\Omega$ ,  $\lambda_{\min}(a)$ , and  $\|(\text{Trace}[\sigma])^{-1}\|_{L^{\infty}(\Omega_T)}$ . If  $n \geq 5$ , it also depends on  $\| \text{Trace}[\sigma] \|_{L^{\infty}(\Omega_T)}$ , and if  $n = 1$ , it also depends on  $\lambda_{\max}(a)$ .

**1.2.2. Medium with a continuum of time scales.**

**THEOREM 1.4.** *Assume that  $\Omega$  is convex, and Condition 1.1 is satisfied; then*

$$(1.30) \quad \| (u - u_h)(T) \|_{L^2(\Omega)} + \| u - u_h \|_{L^2(0, T; H_0^1(\Omega))} \leq Ch \| g \|_{L^2(\Omega_T)}.$$

*Remark 1.11.* The constant  $C$  depends on  $C_X$ ,  $n$ ,  $\Omega$ ,  $\delta$ ,  $\epsilon$ ,  $\lambda_{\min}(a)$ , and  $\lambda_{\max}(a)$ .

The system of ordinary differential equations (1.27) is still characterized by a continuum of time scales in situations where the entries of  $a$  merely belong to  $L^{\infty}(\Omega_T)$ . They need to be discretized (homogenized) in time in order to be solved numerically. This will be the object of the next subsection. Loosely speaking, although (1.1) is associated to a fine tessellation and fine time steps, it is possible to approximate its operator on a coarse tessellation with coarse time steps.

**1.3. Numerical homogenization in space and time.** Let  $M \in \mathbb{N}$ , and  $\{t_n = n \frac{T}{M}\}_{0 \leq n \leq M}$  is a discretization of  $[0, T]$ . Assume  $(\varphi_i)$  is a basis of  $X^h$ . Write  $Z_T^h$ , the subspace of  $Y_T^h$ , such that

$$(1.31) \quad Z_T^h = \left\{ w \in Y_T^h : w(x, t) = \sum_i c_i(t) \varphi_i(F(x, t)), \ c_i(t) \text{ are constants on } (t_n, t_{n+1}] \right\}.$$

Write  $U_T^h$ , the subspace of  $Y_T^h$ , such that

$$(1.32) \quad U_T^h = \left\{ \psi \in Y_T^h : \psi(x, t) = \sum_i d_i \varphi_i(F(x, t)), \ d_i \text{ are constants on } [0, T] \right\}.$$

For  $w \in Y_T^h$ , define  $w_n \in U_T^h$  by

$$(1.33) \quad w_n(x, t) := \sum_i c_i(t_n) \varphi_i(F(x, t)).$$

Write  $v$ , the solution in  $Z_T^h$  of the following system of implicit weak formulation (such that  $v(x, 0) \equiv 0$ ): For  $n \in \{0, \dots, M-1\}$  and  $\psi \in V$ ,

$$(1.34) \quad \begin{aligned} (\psi(t_{n+1}), v_{n+1}(t_{n+1}))_{L^2(\Omega)} &= (\psi(t_n), v_n(t_n))_{L^2(\Omega)} \\ &+ \int_{t_n}^{t_{n+1}} \left( (\partial_t \psi(t), v_{n+1}(t))_{L^2(\Omega)} \right. \\ &\quad \left. - a[\psi(t), v_{n+1}(t)] + (\psi(t), g(t))_{L^2(\Omega)} \right) dt. \end{aligned}$$

The following theorem shows the stability of the implicit scheme (1.34).

**THEOREM 1.5.** *Let  $v \in Z_T$  be the solution of (1.34). We have*

$$(1.35) \quad \|v(T)\|_{L^2(\Omega)} + \|v\|_{L^2(0,T,H_0^1(\Omega))} \leq C \|g\|_{L^2(\Omega_T)}.$$

*Remark 1.12.* The constant  $C$  depends on  $n$ ,  $\Omega$ , and  $\lambda_{\min}(a)$ .

The following theorem gives an error bound on the accuracy of time discretization scheme (1.34) when  $a$  does not depend on time.

**THEOREM 1.6.** *Let  $v \in Z_T$  be the solution of (1.34) and  $u_h$  be the solution of (1.27). Assume that  $\partial_t a \equiv 0$ . We have*

$$(1.36) \quad \begin{aligned} \|(u_h - v)(T)\|_{L^2(\Omega)} + \|u_h - v\|_{L^2(0,T,H_0^1(\Omega))} &\leq C |\Delta t| \\ &\quad \left( \|\partial_t g\|_{L^2(0,T,H^{-1}(\Omega))} + \|g(\cdot, 0)\|_{L^2(\Omega)} \right). \end{aligned}$$

*Remark 1.13.* The constant  $C$  depends on  $n$ ,  $\Omega$ , and  $\lambda_{\min}(a)$ .

*Remark 1.14.* Observe that the combination of Theorem 1.6 with Theorem 1.3 leads to the following inequality:

$$(1.37) \quad \begin{aligned} \|(u - u_h)(\cdot, T)\|_{L^2(\Omega)} + \|u - u_h\|_{L^2(0,T,H_0^1(\Omega))} &\leq C(h + \Delta t) \\ &\quad \left( \|g\|_{L^2(\Omega_T)} + \|\partial_t g\|_{L^2(0,T,H^{-1}(\Omega))} + \|g(\cdot, 0)\|_{L^2(\Omega)} \right). \end{aligned}$$

Hence the accuracy of the method is controlled by  $\max(h, \Delta t)$ .

The following theorem gives an error bound on the accuracy of the time discretization scheme (1.34) when  $a$  has no bounded time derivatives.

**THEOREM 1.7.** *Assume that  $\Omega$  is convex, and Condition 1.1 is satisfied. Let  $v \in Z_T$  be the solution of (1.34) and  $u_h$  be the solution of (1.27), and we have*

$$(1.38) \quad \|(u_h - v)(T)\|_{L^2(\Omega)} + \|u_h - v\|_{L^2(0,T,H_0^1(\Omega))} \leq C \frac{|\Delta t|}{h} \|g\|_{L^2(\Omega_T)},$$

where  $C$  depends on  $\Omega$ ,  $n$ ,  $\delta$ ,  $\epsilon$ ,  $\lambda_{\min}(a)$ , and  $\lambda_{\max}(a)$ .

*Remark 1.15.* Observe that the accuracy of the time discretization scheme (1.34) requires that  $|\Delta t| \ll h$  when  $a$  has no bounded time derivatives. Indeed the combination of Theorem 1.7 with (1.30) leads to the following inequality:

$$(1.39) \quad \begin{aligned} \|(u - u_h)(\cdot, T)\|_{L^2(\Omega)} + \|u - u_h\|_{L^2(0,T,H_0^1(\Omega))} &\leq C \left( h + \frac{\Delta t}{h} \right) \\ &\quad \|g\|_{L^2(\Omega_T)}. \end{aligned}$$

Hence the accuracy of the method is controlled by  $h$  if one chooses  $\Delta t \sim h^2$ . The fact that one must chose  $\Delta t \sim h^2$  for optimal accuracy is a standard phenomenon for parabolic equations, and we refer to [48] for a reminder.

We refer to section 3 for numerical experiments.

**1.4. Remarks on caloric coordinates  $F$ .** In numerical simulations  $a$  is given as mapping from a fine mesh triangulation  $\Omega$  into the set of positive definite symmetric matrices. In other words (in the time-independent case, for instance),  $a$  is a constant matrix in each triangle of a given fine mesh. In the time-dependent case on each fine mesh triangle  $a$  is a step function from  $[0, T]$  into the set of positive definite symmetric matrices.

The caloric coordinates  $F$  have been computed on the same fine triangulation of  $\Omega$  associated to  $a$  through a standard finite element method.

The homogenization results given above assume an underlying continuous medium and that  $F$  is a function on that continuous medium. As has been mentioned in [60] the proofs given in this paper can be extended to the case where the underlying medium is discrete. Indeed the core of those proofs is based on Theorems 2.1 and 2.2, which are straightforward adaptations of Theorem 1.2.1 of [52], and those theorems can be extended to the discrete case.

A similar issue lies in the fact that  $a$  may be a mapping from the continuous medium  $\Omega$ . In that case the equation solved in practice is not (1.1) but a discretization of (1.1) associated to a discretization of  $a$  over a fine mesh triangulation of  $\Omega$ . The control of the error between the solution of (1.1) and the discrete counterpart of (1.1) is standard.

**1.5. Literature and further remarks.** For early works on homogenization with random mixing coefficients we refer to [62], [45], [61], [64], [78], [46], [47], [37]. Papanicolaou, Stroock, and Varadhan [63] have considered a two-component Markov process  $(x(t), y(t))$ , where  $y(t)$  is rapidly varying (and is not assumed to be ergodic in dimension one) and enters in the coefficients of the stochastic process driving  $x(t)$ . They have studied the convergence properties of  $x(t)$  as the fluctuations of  $y(t)$  becomes more rapid using the martingale approach to diffusion, developed by Stroock and Varadhan [67], [66], [69], [68].

Methods dealing with problems involving multiple scales are numerous and not limited to the ones cited in this paper. As an example of a different approach, we refer to [2], where the concept of Young measure on micropatterns is introduced to study variational problems leading to multiple small scales involving a small parameter.

The numerical homogenization method implemented in this paper is a finite element method. The idea of using oscillating test functions can be tracked back to the work of Murat and Tartar on homogenization and  $H$ -convergence; we refer in particular to [71] and [54]. Those papers also contain convergence proofs for the finite element method in an abstract setting for a sequence of  $H$ -converging elliptic operators (recall that the framework of  $H$ -convergence is independent from ergodicity or scale separation assumptions and is based on the compactness of any sequence of solutions of  $-\operatorname{div} a_\epsilon \nabla u_\epsilon = g$ , with uniformly bounded and elliptic conductivities  $a_\epsilon$ ; we also refer to the initial work of Spagnolo [65] for  $G$ -convergence).

The numerical implementation and practical application of oscillating test functions in numerical finite element homogenization have been called multiscale finite element methods and have been studied by several authors [8], [29], [55], [41], [40], [35], [39], [5]. The work of Hou and Wu [41] has been a large source of inspiration in numerical applications (particularly for reservoir modeling in geophysics; we refer to [74], [49], [1], [75] for recent developments) since it was leading to a coarse scale operator while keeping the fine scale structures of the solutions. The construction of the base functions is decoupled from element to element, leading to a scheme adapted to parallel computers. A proof of the convergence of the method is given in periodic

settings when the size of the heterogeneities is smaller than the grid size, and an oversampling technique is proposed to remove the so-called cell resonance error [42] when the size of the heterogeneities is comparable to the grid size.

Allaire and Brizzi [5] have introduced the composition rule in the formulation of a multiscale finite element and observed that a multiscale finite element method with higher order Lagrange polynomials would have a better accuracy. In fact, Babuska et al. introduced a so-called change of variable technique in the general setting of the partition of unity method using a  $p$ -version of finite elements. In [8], a special class of second order elliptic problems with essentially one-dimensional rough coefficients  $a(x, y) = a(x)$  was considered. In [56], [57], conformal mapping was used to map the rough solution to a smoother function for elliptic problems in two dimensions with corners or interfaces.

In [60], we proved that if  $u$  is the solution of the divergence form elliptic equation

$$(1.40) \quad \begin{cases} -\operatorname{div}(a(x)\nabla u(x)) = g & \text{in } \Omega, \\ u = 0 & \text{in } \partial\Omega, \end{cases}$$

and  $F$  are caloric coordinates defined by

$$(1.41) \quad \begin{cases} \operatorname{div} a \nabla F = 0 & \text{in } \Omega, \\ F(x) = x & \text{on } \partial\Omega, \end{cases}$$

then, under the Cordes-type condition  $\beta_\sigma < 1$  on  $\sigma$  given by (1.4), one has for some  $p > 2$

$$(1.42) \quad \|u \circ F^{-1}\|_{W^{2,p}(\Omega)} \leq C \|g\|_{L^p(\Omega)}.$$

It has been deduced from this compensation phenomenon that numerical homogenization methods based on oscillating finite elements can converge in the presence of a continuum of scales if one uses global caloric coordinates to obtain the test functions instead of solutions of a local cell problem [60]. In dimension three and higher it has been known since the work of Fenchenco and Khruslov [34], [44] that the homogenization of divergence form elliptic operators  $-\operatorname{div} a_\epsilon \nabla u_\epsilon = g$  can lead to a nonlocal homogenized operator if the sequence of matrices  $a_\epsilon$  is uniformly elliptic but with entries uniformly bounded only in  $L^1(\Omega)$ . From a numerical point of view these nonlocal effects imply that a nonlocal numerical homogenization method cannot be avoided to obtain accuracy. Hence in [60], it is shown that the accuracy of local methods depends on the aspect ratio of the triangles of the tessellation with respect to caloric coordinates (which is not the case if one uses nonlocal finite elements; we refer to [60] for further discussions on the apparition of nonlocal effects in numerical homogenization). Recently Briane has shown [21] that this nonlocal effect is absent in dimension two in the  $H$ -convergence setting.

The phenomenon is similar here; however, observe that, if one has solved the initial parabolic equation at least  $n$  times and those solutions are (locally) linearly independent, it is also possible to use them as new coordinates for numerical homogenization. Observe that in dimension higher than three the caloric coordinates are not always invertible; an idea to bypass this difficulty could be either to choose the change of coordinates locally and adaptively or to enrich the coordinates by writing down the initial equations as degenerate equations in a space of higher dimension [72]; these points have not been explored. For divergence form elliptic equations, recall that fast methods based on hierarchical matrices are available [14], [13], [12], [10], [11] for solving

(1.40) and (1.41) in  $O(N(\ln N)^{n+3})$  operations (or in  $O(N)$  operations using iterative methods (see [76], [77]),  $N$  being the number of interior nodes of the fine mesh).

The issue of numerical homogenization partial differential equations with heterogeneous coefficients has received a great deal of attention, and many methods have been proposed. A few of them are cited below.

- multiscale finite volume methods [50];
- heterogeneous multiscale methods [28], [24];
- wavelet-based homogenization [38], [27], [25], [17], [7], [19];
- residual-free bubble methods [20];
- discontinuous enrichment methods [33], [32];
- partition of unity methods [36];
- energy-minimizing multigrid methods [73].

Following [60], it is possible to implement a finite volume method based on the compensation theorems given in this paper. The elements given in this paper contain the fine scale structure of  $F$ , and has been done in [60], and it is possible to approximate the initial parabolic operator by a homogenized parabolic operator associated to the coarse mesh (the test functions in this case would be piecewise linear on the coarse mesh, and the approximation error associated to the homogenized operator would depend on the aspect ratio of the triangles of the coarse mesh in the metric induced by  $F$ ).

In this paper  $a$  has been assumed to be bounded and uniformly elliptic. Without these assumptions the diffusion associated to a homogenized operator can be anomalously slow [15], [58] or fast (superdiffusive) [59]. If  $a$  has an unbounded skew symmetric component, the homogenization of (1.1) can give rise to a degenerate operator [59].

## 2. Proofs.

### 2.1. Compensation.

**2.1.1. Time-independent medium: Proof of Theorem 1.1.** We will need the following lemmas. Let  $\mathcal{A}_T$  be the bilinear form on  $L^2(0, T; H_0^1(\Omega))$  defined by

$$(2.1) \quad \mathcal{A}_T[v, w] := \int_0^T a[v, w](t) dt,$$

where

$$(2.2) \quad a[v, u](t) := \int_{\Omega} {}^t \nabla v(x, t) a(x, t) \nabla u(x, t) dx.$$

We write  $\mathcal{A}_T[u] := \mathcal{A}_T[u, u]$ . The following lemmas are standard energy estimates, and we refer to [31, section 7.1.2] for a reminder.

LEMMA 2.1. *We have*

$$(2.3) \quad \|u(\cdot, T)\|_{L^2(\Omega)}^2 + \mathcal{A}_T[u] \leq \frac{C_{n, \Omega}}{\lambda_{\min}(a)} \|g\|_{L^2(\Omega_T)}^2.$$

LEMMA 2.2. *Assume  $\partial_t a \equiv 0$ , and we have*

$$(2.4) \quad \|\partial_t u\|_{L^2(\Omega_T)}^2 + a[u(\cdot, T)] \leq \|g\|_{L^2(\Omega_T)}^2.$$

LEMMA 2.3. *Assume  $\partial_t a \equiv 0$ , and we have*

$$(2.5) \quad \|\partial_t u(\cdot, T)\|_{L^2(\Omega)}^2 + \mathcal{A}_T[\partial_t u] \leq \frac{C_{n, \Omega}}{\lambda_{\min}(a)} \|\partial_t g\|_{L^2(0, T, H^{-1}(\Omega))}^2 + \|g(\cdot, 0)\|_{L^2(\Omega)}^2.$$

We now need a variation of Campanato's result [23] on nondivergence form elliptic operators. Let us write for a symmetric matrix  $M$

$$(2.6) \quad \nu_M := \frac{\text{Trace}(M)}{\text{Trace}({}^t M M)}.$$

We consider the following Dirichlet problem:

$$(2.7) \quad L_M v = f,$$

with  $L_M := \sum_{i,j=1}^n M_{ij}(x) \partial_i \partial_j$ . The following Theorems 2.1 and 2.2 are straightforward adaptations of Theorem 1.2.1 of [52]. They are proven in [52] under the assumption that  $M$  is bounded and elliptic. It is easy to check that the conditions  $\beta_M < 1$  and  $\nu_M \in L^\infty(\Omega)$  are sufficient for the validity of those theorems. We refer to [60] for that adaptation.

**THEOREM 2.1.** *Assume that  $\beta_M < 1$ ,  $\nu_M \in L^\infty(\Omega)$ , and  $\Omega$  is convex. Then if  $f \in L^2(\Omega)$ , the Dirichlet problem (2.7) has a unique solution satisfying*

$$(2.8) \quad \|v\|_{W_D^{2,2}(\Omega)} \leq \frac{C}{1 - \beta_M^{\frac{1}{2}}} \|\nu_M f\|_{L^2(\Omega)}.$$

*Remark 2.1.*  $\beta_M$  is the Cordes parameter associated to  $M$ .

**THEOREM 2.2.** *Assume that  $\beta_M < 1$ ,  $\nu_M \in L^\infty(\Omega)$ , and  $\Omega$  is convex. Then there exists a real number  $p_0 > 2$  depending only on  $n, \Omega$ , and  $\beta_M$  such that for each  $f \in L^p(\Omega)$ ,  $2 \leq p < p_0$ , the Dirichlet problem (2.7) has a unique solution satisfying*

$$(2.9) \quad \|v\|_{W_D^{2,p}(\Omega)} \leq \frac{C_{n,\Omega,p}}{1 - \beta_M^{\frac{1}{2}}} \|\nu_M f\|_{L^p(\Omega)}.$$

Let us now prove the compensation theorems. Choose

$$(2.10) \quad M := \frac{\sigma}{|\det(\nabla F)|^{\frac{1}{2}}} \circ F^{-1}.$$

It is easy to check that  $\beta_\sigma < 1$  implies that  $F$  is a homeomorphism from  $\Omega$  onto  $\Omega$ , and thus (2.10) is well defined. Moreover observe that  $\beta_M = \beta_\sigma$  and

$$(2.11) \quad \|\nu_M\|_{L^\infty(\Omega_T)}^2 \leq \frac{C_n}{(\lambda_{\min}(a))^{\frac{n}{2}}} \|(\text{Trace}[\sigma])^{\frac{n}{4}-1}\|_{L^\infty(\Omega_T)}^2.$$

Fix  $t \in [0, T]$ . Choose

$$(2.12) \quad f := \frac{(\partial_t u - g)}{|\det(\nabla F)|^{\frac{1}{2}}} \circ F^{-1}.$$

Observe that by the change of variable  $y = F(x)$  one obtains that, if  $\partial_t a \equiv 0$  (which implies that  $F$  is time-independent),  $\partial_t u \in L^2(\Omega)$ , and  $g(\cdot, t) \in L^2(\Omega)$ ,  $f \in L^2(\Omega)$  and

$$(2.13) \quad \|f\|_{L^2(\Omega)} \leq \|\partial_t u\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}.$$

It follows from Theorem 2.1 that there exists a unique  $v \in W_D^{2,2}(\Omega)$  satisfying

$$(2.14) \quad \|v\|_{W_D^{2,2}(\Omega)}^2 \leq \frac{C \|\nu_M\|_{L^\infty(\Omega_T)}^2}{(1 - \beta_\sigma^{\frac{1}{2}})^2} (\|\partial_t u\|_{L^2(\Omega)}^2 + \|g\|_{L^2(\Omega)}^2)$$

and the following equation:

$$(2.15) \quad \partial_t \hat{u}(y, t) = \sum_{i,j} (\sigma(F^{-1}(y, t), t))_{i,j} \partial_i \partial_j v(y, t) + \hat{g}(y, t).$$

We use the notation  $\hat{g} := g \circ F^{-1}$  and  $\hat{u} := u \circ F^{-1}$ . Using the change of variable  $y = F(x)$  and using the property  $\operatorname{div} a \nabla F = 0$  when  $\partial_t a \equiv 0$ , we obtain that (2.15) can be written

$$(2.16) \quad \partial_t u = \operatorname{div} (a \nabla (v \circ F)) + g.$$

If  $\partial_t u \in L^2(\Omega)$  and  $g(\cdot, t) \in L^2(\Omega)$ , we can use the uniqueness property of the solution of the divergence form elliptic Dirichlet problem

$$(2.17) \quad \operatorname{div} (a \nabla w) = \partial_t u - g$$

to obtain that  $v \circ F = u$ . Assume that  $g \in L^2(\Omega_T)$  and  $\partial_t u \in L^2(\Omega_T)$ . It follows that, for  $t \in [0, T] - B$ ,  $g(\cdot, t) \in L^2(\Omega)$  and  $\partial_t u(\cdot, t) \in L^2(\Omega)$ , where  $B$  is a subset of  $[0, T]$  of 0-Lebesgue measure. It follows from the previous arguments that, on  $[0, T] - B$ ,  $u \circ F^{-1}(\cdot, t) \in W_D^{2,2}(\Omega)$  and satisfies

$$(2.18) \quad \|u \circ F^{-1}(\cdot, t)\|_{W_D^{2,2}(\Omega)}^2 \leq \frac{C \|\nu_M\|_{L^\infty(\Omega_T)}^2}{(1 - \beta_\sigma^{\frac{1}{2}})^2} (\|\partial_t u(\cdot, t)\|_{L^2(\Omega)}^2 + \|g(\cdot, t)\|_{L^2(\Omega)}^2).$$

Integrating (2.18) with respect to time we obtain that  $u \circ F^{-1} \in L^2(0, T, W_D^{2,2}(\Omega))$  and that

$$(2.19) \quad \|u \circ F^{-1}\|_{L^2(0, T, W_D^{2,2}(\Omega))}^2 \leq \frac{C \|\nu_M\|_{L^\infty(\Omega_T)}^2}{(1 - \beta_\sigma^{\frac{1}{2}})^2} (\|\partial_t u\|_{L^2(\Omega_T)}^2 + \|g\|_{L^2(\Omega_T)}^2).$$

Thus using Lemma 2.2 we have obtained Theorem 1.1.

In situations where  $g \in L^\infty(0, T, L^2(\Omega))$ ,  $\partial_t g \in L^2(0, T, H^{-1}(\Omega))$ , or  $g \in L^p(\Omega_T)$ , with  $p > 2$ , one can obtain a higher regularity for  $u \circ F^{-1}$ . This is the object of the following theorems.

**THEOREM 2.3.** *Assume that  $\Omega$  is convex,  $g \in L^\infty(0, T, L^2(\Omega))$ ,  $\partial_t g \in L^2(0, T, H^{-1}(\Omega))$ ,  $\partial_t a \equiv 0$ ,  $\beta_\sigma < 1$ , and  $(\operatorname{Trace}[\sigma])^{\frac{n}{4}-1} \in L^\infty(\Omega_T)$ ; then for all  $t \in [0, T]$ ,  $u \circ F^{-1}(\cdot, t) \in W_D^{2,2}(\Omega)$  and*

$$(2.20) \quad \|u \circ F^{-1}(\cdot, t)\|_{W_D^{2,2}(\Omega)} \leq \frac{C}{1 - \beta_\sigma^{\frac{1}{2}}} \left( \|g\|_{L^\infty(0, T, L^2(\Omega))} + \|\partial_t g\|_{L^2(0, T, H^{-1}(\Omega))} \right).$$

*Remark 2.2.* The constant  $C$  can be written

$$C = \frac{C_{n,\Omega}}{(\lambda_{\min}(a))^{\frac{n}{4}}} \|(\operatorname{Trace}[\sigma])^{\frac{n}{4}-1}\|_{L^\infty(\Omega_T)} \left( 1 + \frac{1}{\lambda_{\min}(a)} \right)^{\frac{1}{2}}.$$

**THEOREM 2.4.** *Assume that  $\Omega$  is convex,  $g(\cdot, 0) \in L^2(\Omega)$ ,  $\partial_t g \in L^2(0, T, H^{-1}(\Omega))$ ,  $g \in L^p(\Omega_T)$ ,  $\partial_t a \equiv 0$ ,  $\beta_\sigma < 1$ , and  $(\operatorname{Trace}[\sigma])^{\frac{n}{4}-1} \in L^\infty(\Omega_T)$ ; then there exists a real number  $p_0 > 2$  depending only on  $n, \Omega$ , and  $\beta_\sigma$  such that for each  $p$  such that  $2 \leq p < p_0$  one has*

$$(2.21) \quad \|u \circ F^{-1}\|_{L^p(0, T, W_D^{2,p}(\Omega))} \leq \frac{C}{1 - \beta_\sigma^{\frac{1}{2}}} (\|g\|_{L^p(\Omega_T)} + \|g(\cdot, 0)\|_{L^2(\Omega)} + \|\partial_t g\|_{L^2(0, T, H^{-1}(\Omega))}).$$

*Remark 2.3.* The constant  $C$  can be written

$$C = \frac{C_{n,\Omega,p}}{(\lambda_{\min}(a))^{\frac{n}{4}}} \|(\text{Trace}[\sigma])^{\frac{n}{4}-1}\|_{L^\infty(\Omega_T)} \left(1 + \frac{1}{\lambda_{\min}(a)}\right)^{\frac{1}{2}}.$$

Write

$$(2.22) \quad \|v\|_{C^\gamma(\Omega)} := \sup_{x,y \in \Omega, x \neq y} \frac{|v(x) - v(y)|}{|x - y|^\gamma}.$$

**THEOREM 2.5.** *Assume that  $n \leq 2$ ,  $\Omega$  is convex,  $g(\cdot, 0) \in L^2(\Omega)$ ,  $\partial_t g \in L^2(0, T, H^{-1}(\Omega))$ ,  $g \in L^p(\Omega_T)$ ,  $\partial_t a \equiv 0$ ,  $\beta_\sigma < 1$ ,  $(\text{Trace}[\sigma])^{-1} \in L^\infty(\Omega_T)$ , and  $g \in L^2[0, T; L^{p^*}(\Omega)]$ , with  $2 < p^*$ . Then there exists  $p \in (2, p^*]$  and  $\gamma(p) > 0$  such that*

$$(2.23) \quad \left( \int_0^T \|\nabla(u \circ F^{-1})(\cdot, t)\|_{C^\gamma(\Omega)}^2 dt \right)^{\frac{1}{2}} \leq \frac{C}{1 - \beta_\sigma^{\frac{1}{2}}} (\|g\|_{L^p(\Omega_T)} + \|g(\cdot, 0)\|_{L^2(\Omega)} + \|\partial_t g\|_{L^2(0, T, H^{-1}(\Omega))}).$$

*Remark 2.4.* The constant  $C$  in (2.23) depends on  $n$ ,  $p$ ,  $\Omega$ ,  $\lambda_{\min}(a)$ , and  $\|(\text{Trace}(\sigma))^{-1}\|_{L^\infty(\Omega_T)}$ . It is easy to check that if  $n = 1$ , then the theorem is valid with  $\gamma = 1/2$ .

Using Lemma 2.3 we can prove Theorem 2.3.

Let us now prove Theorem 2.4. Assume that there exists  $q_0 > 2$  such that, for  $2 \leq p < q_0$ ,  $\partial_t u \in L^p(\Omega_T)$  and  $g \in L^p(\Omega_T)$ . Let us now apply Theorem 2.2 with  $p < \min(p_0, q_0)$ ,  $M$  given by (2.10), and  $f$  given by (2.12). It follows that, for  $t \in [0, T] - B$  (where  $B$  is a subset of  $[0, T]$  of 0-Lebesgue measure),  $g(\cdot, t) \in L^p(\Omega)$  and  $\partial_t u(\cdot, t) \in L^p(\Omega)$ . We deduce from Theorem 2.2 and the argumentation related to (2.17) that, on  $[0, T] - B$ ,  $u \circ F^{-1}(\cdot, t) \in W_D^{2,p}(\Omega)$  and

$$(2.24) \quad \|u \circ F^{-1}(\cdot, t)\|_{W_D^{2,p}(\Omega)}^p \leq \frac{C_{n,p,\Omega} \|\nu_M\|_{L^\infty(\Omega_T)}^p}{(1 - \beta_\sigma^{\frac{1}{2}})^p} (\|\partial_t u(\cdot, t)\|_{L^p(\Omega)}^p + \|g(\cdot, t)\|_{L^p(\Omega)}^p).$$

Integrating (2.24) with respect to time, we have  $u \circ F^{-1} \in L^p(0, T, W_D^{2,p}(\Omega))$  and

$$(2.25) \quad \|u \circ F^{-1}\|_{L^p(0, T, W_D^{2,p}(\Omega))} \leq \frac{C_{n,p,\Omega} \|\nu_M\|_{L^\infty(\Omega_T)}}{1 - \beta_\sigma^{\frac{1}{2}}} (\|\partial_t u\|_{L^p(\Omega_T)} + \|g\|_{L^p(\Omega_T)}).$$

It remains to show that, under the assumptions of Theorem 2.4,  $\partial_t u \in L^p(\Omega_T)$ .

In order to bound  $\|\partial_t u(\cdot, t)\|_{L^p(\Omega)}$  we use general Sobolev inequalities (chapter 5.6 of [31]).

- If  $n \geq 3$ , write  $p^* = 2n/(n-2)$ . We have for  $2 < p \leq p^*$

$$(2.26) \quad \left( \int_\Omega (\partial_t u)^p dx \right)^{\frac{2}{p}} \leq C_{n,\Omega} \left( \int_\Omega (\partial_t u)^{p^*} dx \right)^{\frac{2}{p^*}},$$

and thus, using the Gagliardo–Nirenberg–Sobolev inequality,

$$(2.27) \quad \left( \int_\Omega (\partial_t u)^p dx \right)^{\frac{2}{p}} \leq C_{n,p,\Omega} \frac{1}{\lambda_{\min}(a)} a[\partial_t u].$$



- If  $n = 2$ , we write, for  $(x_1, x_2, x_3) \in \Omega \times (0, 1)$ ,  $v(x_1, x_2, x_3) := \partial_t u(x_1, x_2)$ . Using the Gagliardo–Nirenberg–Sobolev inequality in dimension three we obtain that for  $2 < p \leq 6$

$$(2.28) \quad \left( \int_{\Omega} (\partial_t u)^p dx \right)^{\frac{2}{p}} \leq C_{n,p,\Omega} \int_{\Omega} (\nabla \partial_t u)^2 dx,$$

which leads us to (2.27).

- If  $n = 1$ , then using Morrey's inequality we obtain that, with  $\gamma := 1/2$ ,

$$(2.29) \quad \|\partial_t u\|_{C^{0,\gamma}(\Omega)}^2 \leq C_{\Omega} \frac{1}{\lambda_{\min}(a)} a[\partial_t u].$$

We conclude the proof of Theorem 2.4 by using Lemma 2.3.

We deduce Theorem 2.5 from Morrey's inequality and Theorem 2.4.

*Hölder continuity for  $n \geq 3$  or nonconvexity of  $\Omega$ .* In this paragraph we will not assume  $\Omega$  to be convex. Let  $N^{p,\lambda}(\Omega)$  ( $1 < p < \infty$ ,  $0 < \lambda < n$ ) be the weighted Morrey space formed by the functions  $v : \Omega \rightarrow \mathbb{R}$  such that  $\|v\|_{N^{p,\lambda}(\Omega)} < \infty$ , with

$$(2.30) \quad \|v\|_{N^{p,\lambda}(\Omega)} := \sup_{x_0 \in \Omega} \left( \int_{\Omega} |x - x_0|^{-\lambda} |v(x)|^p dx \right)^{\frac{1}{p}}.$$

To obtain the Hölder continuity of  $u \circ F^{-1}$  in dimension  $n \geq 3$  we use Corollary 4.1 of [51]. We give the result of Leonardi below in a form adapted to our context. Consider the Dirichlet problem (2.7). We write  $W^{2,p,\lambda}(\Omega)$ , the functions in  $W_D^{2,p}(\Omega)$ , such that their second order derivatives belong to  $N^{p,\lambda}(\Omega)$ .

**THEOREM 2.6.** *There exist a constant  $C^* = C^*(n, p, \lambda, \partial\Omega) > 0$  such that if  $\beta_M < C^*$  and  $f \in N^{p,\lambda}(\Omega)$ , then the Dirichlet problem (2.7) has a unique solution in  $W^{2,p,\lambda} \cap W_0^{1,p}(\Omega)$ . Moreover, if  $0 < \lambda < n < p$ , then  $\nabla v \in C^{\alpha}(\Omega)$ , with  $\alpha = 1 - n/p$  and*

$$(2.31) \quad \|\nabla v\|_{C^{\alpha}(\Omega)} \leq \frac{C}{\lambda_{\min}(M)} \|f\|_{N^{p,\lambda}(\Omega)},$$

where  $C = C(n, p, \lambda, \partial\Omega)$ .

Now we have the compensation theorem in the nonconvex case.

**THEOREM 2.7.** *Assume  $n \geq 2$  and  $\partial_t a \equiv 0$ . Let  $p > 2$ . There exists a constant  $C^* = C^*(n, \partial\Omega) > 0$ , a real number  $\gamma > 0$ , depending only on  $n, \Omega$ , and  $p$  such that if  $\beta_{\sigma} < C^*$ , then*

$$(2.32) \quad \left( \int_0^T \|\nabla(u \circ F^{-1})(\cdot, t)\|_{C^{\gamma}(\Omega)}^2 dt \right)^{\frac{1}{2}} \leq C(\|g\|_{L^p(\Omega_T)} + \|g(\cdot, 0)\|_{L^2(\Omega)} + \|\partial_t g\|_{L^2(0,T,H^{-1}(\Omega))}).$$

**Remark 2.5.** The constant  $C$  in (2.32) depends on  $n, \gamma, \Omega, C^*, \lambda_{\min}(a)$ , and  $\|(\text{Trace}(\sigma))^{\frac{n}{2p}-1}\|_{L^{\infty}(\Omega_T)}$ .

**Remark 2.6.** Compared with the convex case, we have to assume  $F$  is an automorphism. Theorem 2.7 shows that convexity is not needed, but without convexity the Cordes-type assumption  $\beta_{\sigma} < C^*$  is stronger than  $\beta_{\sigma} < 1$ .

The proof of Theorem 2.7 is an application of Theorem 2.6. We just need to observe that from the Hölder inequality we have for  $0 < \epsilon < 0.5$

$$(2.33) \quad \|f\|_{N^{p,\epsilon}(\Omega)} \leq C_{n,p,\Omega,\epsilon} \|f\|_{L^{p(1+\epsilon)}(\Omega)}.$$

From this point the proof is similar to the proof of Theorem 2.4.

**2.1.2. Medium with a continuum of time scales: Proof of Theorem 1.2.**

We will need Theorems 1.6.2 and 1.6.3 of [52]. For the sake of completeness we will recall those theorems below in a version adapted to our framework. Consider the following parabolic problem:

$$(2.34) \quad \partial_t v = \sum_{i,j=1}^n M_{ij}(x) \partial_i \partial_j v + f.$$

We assume  $M$  to be symmetric bounded and elliptic and  $v = 0$  at  $t = 0$  and on the boundary  $\partial\Omega$ . Write

$$(2.35) \quad \eta_M := \sup_{x \in \Omega_T} \frac{\text{Trace}[{}^t M M] + 1}{(\text{Trace}[M] + 1)^2}$$

and

$$(2.36) \quad \alpha_M := \sup_{x \in \Omega_T} \frac{\text{Trace}[M] + 1}{\text{Trace}[{}^t M M] + 1}.$$

Write for  $p \geq 2$

$$(2.37) \quad S_p(\Omega_T) := \left\{ v \in L^p(0, T, W_D^{2,p}(\Omega)); \partial_t v \in L^p(\Omega_T); v(\cdot, 0) \equiv 0 \right\}$$

and

$$(2.38) \quad \|v\|_{S_p(\Omega_T)}^p := \int_{\Omega_T} \left( \sum_{i,j} (\partial_i \partial_j v)^2 + (\partial_t v)^2 \right)^{\frac{p}{2}} dy dt.$$

**THEOREM 2.8.** *Assume  $\Omega$  to be convex and that there exists  $\epsilon \in (0, 1)$  such that  $\eta_M \leq 1/(n+\epsilon)$ ; then for each  $f \in L^2(\Omega_T)$  the Cauchy–Dirichlet problem (2.34) admits a unique solution in  $S_2(\Omega_T)$  which satisfies the bound*

$$(2.39) \quad \|v\|_{S_2(\Omega_T)} \leq \frac{\alpha_M}{1 - \sqrt{1 - \epsilon}} \|f\|_{L^2(\Omega_T)}.$$

**THEOREM 2.9.** *Assume  $\Omega$  to be convex and that there exists  $\epsilon \in (0, 1)$  such that  $\eta_M \leq 1/(n+\epsilon)$ ; then there exists a number  $p_0 > 2$  depending on  $\Omega, n, \epsilon$  such that for each  $f \in L^p(\Omega_T)$  the Cauchy–Dirichlet problem (2.34) admits a unique solution in  $S_p(\Omega_T)$  which satisfies the bound*

$$(2.40) \quad \|v\|_{S_p(\Omega_T)} \leq C_p \frac{\alpha_M}{1 - \sqrt{1 - \epsilon}} \|f\|_{L^p(\Omega_T)}.$$

*Remark 2.7.* In fact Theorem 1.6.3 of [52] is written with  $1 - C(p)\sqrt{1 - \epsilon}$  in the denominator of (2.40), but it is easy to modify it to obtain (2.40) by lowering the value of  $p_0$  and changing the value of  $C_p$ .

Let  $\delta > 0$ . Let us now prove Proposition 1.1. Write  $x = \text{Trace}[\sigma]$  and  $z = n \frac{\text{Trace}[{}^t \sigma \sigma]}{(\text{Trace}[\sigma])^2}$  (observe that  $1 \leq z \leq n$ ). It is easy to check that Condition 1.1 can be written

$$(2.41) \quad -\delta^2 x^2 \left( \frac{\epsilon + n}{n} z - 1 \right) + 2x\delta - (n + \epsilon - 1) \geq 0.$$

Choose  $\delta = n \|(\text{Trace}[\sigma])^{-1}\|_{L^\infty(\Omega_T)}$ . Observing that  $\delta x \geq n$  and  $\delta x = ny_\sigma$  it is easy to conclude the proof of Proposition 1.1. Similarly one obtains the following lemma by straightforward computation from (2.41).

LEMMA 2.4. *Assume that Condition 1.1 is satisfied; then  $\mu_\sigma < C(n, \epsilon, \delta)$*

$$(2.42) \quad \|(\text{Trace}[\sigma])^{-1}\|_{L^\infty(\Omega_T)} \leq C(n, \epsilon, \delta)$$

and

$$(2.43) \quad \|\text{Trace}[\sigma]\|_{L^\infty(\Omega_T)} \leq C(n, \epsilon, \delta).$$

Let us now apply Theorem 2.8 on  $[0, T/\delta]$  with

$$(2.44) \quad M := \delta \sigma \circ F^{-1}(y, \delta t)$$

and

$$(2.45) \quad f := \delta(g \circ F^{-1})(y, \delta t).$$

Observe that if Condition 1.1 is satisfied, then  $F$  is a homeomorphism and  $M$  is well-defined, bounded, and elliptic. Moreover  $\eta_M < \infty$  and  $\alpha_M < \infty$  since

$$(2.46) \quad \text{esssup}_{\Omega_T} \frac{\text{Trace}[^t M M] + 1}{(\text{Trace}[M] + 1)^2} = \text{esssup}_{\Omega_T} \frac{\delta^2 \text{Trace}[^t \sigma \sigma] + 1}{(\delta \text{Trace}[\sigma] + 1)^2}.$$

It follows that the following equation admits a unique solution in  $S_2(\Omega_{\frac{T}{\delta}})$ :

$$(2.47) \quad \partial_t w(y, t) = \sum_{i,j} M_{i,j}(y, t) \partial_i \partial_j w(y, t) + k(y, t),$$

with  $k(y, t) = \delta \hat{g}(y, \delta t)$ . We also have

$$(2.48) \quad \int_0^{\frac{T}{\delta}} \int_{\Omega} \left( (\partial_t w)^2 + \sum_{i,j} (\partial_i \partial_j w)^2 \right) dy dt \leq \frac{C}{(1 - \sqrt{1 - \epsilon})^2} \|f\|_{L^2(\Omega_{\frac{T}{\delta}})}.$$

Using the change of variables  $t \rightarrow \delta t$  and writing

$$(2.49) \quad w(y, t) := v(y, \delta t),$$

we obtain that  $v$  satisfies the following equation on  $\Omega_T$ :

$$(2.50) \quad \partial_t v(y, t) = \sum_{i,j} (\sigma(F^{-1}(y, t), t))_{i,j} \partial_i \partial_j v(y, t) + \hat{g}(y, t).$$

Using the change of variable  $y = F(x)$  and observing that  $\partial_t F = \text{div } a \nabla F$ , we obtain that  $v \circ F$  satisfies

$$(2.51) \quad \partial_t (v \circ F) = \text{div} (a \nabla (v \circ F)) + g.$$

It follows from the uniqueness of the solution of (2.51) that  $u = v \circ F$ . In resume we have obtained Theorem 1.2 (we use Lemma 2.4 to control the constants).

Similarly we obtain the following theorems in situations where  $g \in L^p(\Omega_T)$ , with  $p > 2$ .

**THEOREM 2.10.** *Assume that  $\Omega$  is convex and Condition 1.1 is satisfied; then there exists a number  $p_0 > 2$  depending on  $n, \Omega, \epsilon$  such that, for  $p \in (2, p_0)$ ,  $u \circ F^{-1} \in L^p(0, T, W_D^{2,p}(\Omega))$ ,  $\partial_t(u \circ F^{-1}) \in L^p(\Omega_T)$ , and*

$$(2.52) \quad \|u \circ F^{-1}\|_{L^p(0,T,W_D^{2,p}(\Omega))} + \|\partial_t(u \circ F^{-1})\|_{L^p(\Omega_T)} \leq C\|g\|_{L^p(\Omega_T)},$$

where  $C$  depends on  $\Omega$ ,  $n$ ,  $\delta$ , and  $\epsilon$ .

**THEOREM 2.11.** *Assume that  $\Omega$  is convex and Condition 1.1 is satisfied; then there exists a number  $\alpha_0 > 2$  depending on  $n, \Omega, \epsilon, p$  such that, for  $\alpha \in (0, \alpha_0)$ ,  $\nabla(u \circ F^{-1}) \in L^2(0, T, C^\alpha(\Omega))$  and*

$$(2.53) \quad \|\nabla(u \circ F^{-1})\|_{L^2(0,T,C^\alpha(\Omega))} \leq C\|g\|_{L^p(\Omega_T)},$$

where  $C$  depends on  $\Omega$ ,  $\delta$ ,  $n$ , and  $\epsilon$ .

The proof of Theorem 2.10 is similar and based on Theorem 2.9. The proof of Theorem 2.11 follows from Theorem 2.10 and Morrey's inequality.

**2.2. Convergence of the numerical homogenization.** Write  $\mathcal{R}_h$ , the projection operator mapping  $L^2(0, T; H_0^1(\Omega))$  onto  $Y_T^h$  defined by: For all  $v \in Y_T^h$

$$(2.54) \quad \mathcal{A}_T[v, u - \mathcal{R}_h u] = 0.$$

Write  $\rho := u - \mathcal{R}_h u$  and  $\theta := \mathcal{R}_h u - u_h$ . By a standard energy estimate, we have the following lemmas without proof.

**LEMMA 2.5.**

$$(2.55) \quad \frac{1}{2} \|(u - u_h)(T)\|_{L^2(\Omega)}^2 + \mathcal{A}_T[u - u_h] = \int_{\Omega_T} \rho \partial_t(u - u_h) + \mathcal{A}_T[\rho, u - u_h].$$

*Proof.* Subtracting (1.1) (integrated against  $\psi$ ) and (1.27) we obtain that

$$(2.56) \quad (\psi, \partial_t(u - u_h)) + a[\psi, u - u_h] = 0 \quad \text{for all } \psi \in V_h(t).$$

Integrating by parts with respect to time we deduce that

$$(2.57) \quad (\psi, (u - u_h)(\cdot, t)) + a[\psi, u - u_h] = \int_{\Omega_t} \partial_t \psi (u - u_h).$$

Taking  $\psi = \theta$  in (2.57) we deduce that

$$(2.58) \quad \|(u - u_h)(\cdot, t)\|_{L^2(\Omega)}^2 + \mathcal{A}_t[u - u_h] = \int_{\Omega_t} \partial_t \theta (u - u_h) + (\rho, (u - u_h)(\cdot, t)) + \mathcal{A}_t[\rho, u - u_h].$$

Observing that

$$(2.59) \quad \int_0^t (\partial_t \theta, u - u_h) + (\rho, (u - u_h)(\cdot, t)) = \frac{1}{2} \|(u - u_h)(\cdot, t)\|_{L^2(\Omega)}^2 + \int_0^t (\rho, \partial_t(u - u_h))$$

we deduce the lemma.  $\square$

**2.2.1. Time-independent medium: Proof of Theorem 1.3.**

LEMMA 2.6.

$$(2.60) \quad \|(u - u_h)(T)\|_{L^2(\Omega)}^2 + \mathcal{A}_T[u - u_h] \leq 2 \left( \|\rho\|_{L^2(\Omega_T)} \|\partial_t u - \partial_t u_h\|_{L^2(\Omega_T)} + \mathcal{A}_T[\rho] \right).$$

*Proof.* Lemma 2.6 is a straightforward consequence of Lemma 2.5 and Cauchy–Schwartz and Minkowski inequalities.  $\square$

Similar to Lemmas 2.1 and 2.2, we have the following energy estimate for  $u_h$ .

LEMMA 2.7. Assume  $\partial_t a \equiv 0$ , and we have

$$(2.61) \quad \|u_h(\cdot, T)\|_{L^2(\Omega)}^2 + \mathcal{A}_T[u_h] \leq \frac{C_{n,\Omega}}{\lambda_{\min}(a)} \|g\|_{L^2(\Omega_T)}^2.$$

LEMMA 2.8. Assume  $\partial_t a \equiv 0$ , and we have

$$(2.62) \quad \|\partial_t u_h\|_{L^2(\Omega_T)}^2 + a[u_h(\cdot, T)] \leq \|g\|_{L^2(\Omega_T)}^2.$$

Let  $t \in [0, T]$  and  $v \in H_0^1(\Omega)$ , and we will write  $\mathcal{R}_{h,t}v(\cdot, t)$ , the solution of

$$(2.63) \quad \int_{\Omega} {}^t\nabla \psi a(x, t) (\psi, v - \mathcal{R}_{h,t}v) dx = 0 \quad \text{for all } \psi \in V_h(t).$$

We will need the following lemma.

LEMMA 2.9. Assume the mapping  $x \rightarrow F(x, t)$  to be invertible, and then for  $v \in H_0^1(\Omega)$  we have

- for  $n = 1$ ,

$$(2.64) \quad (a[v - \mathcal{R}_{h,t}v])^{\frac{1}{2}} \leq C_X h \|v \circ F^{-1}(\cdot, t)\|_{W_D^{2,2}} \|a \nabla F\|_{L^\infty(\Omega_T)}^{\frac{1}{2}};$$

- for  $n \geq 2$ ,

$$(2.65) \quad (a[v - \mathcal{R}_{h,t}v])^{\frac{1}{2}} \leq C_X h \|v \circ F^{-1}(\cdot, t)\|_{W_D^{2,2}} \times C_n \mu_\sigma^{\frac{n-1}{4}} \|(\text{Trace}[\sigma])^{-1}\|_{L^\infty(\Omega_T)}^{\frac{n-2}{4}}.$$

*Remark 2.8.* Recall that  $\mu_\sigma$  is given by (1.12), and it is easy to check that  $\mu_\sigma$  is bounded by an increasing function of  $(1 - \beta_\sigma)^{-1}$ .

*Proof.* Using the change of coordinates  $y = F(x, t)$  we obtain that (we write  $\hat{v} := v \circ F^{-1}$ )

$$(2.66) \quad a[v] = Q[\hat{v}],$$

with

$$(2.67) \quad Q[w] := \int_{\Omega} {}^t\nabla w(y, t) Q(y, t) \nabla w(y, t) dy$$

and

$$(2.68) \quad Q(y, t) := \frac{\sigma}{\det(\nabla F)} \circ F^{-1}.$$

Using the definition of  $\mathcal{R}_{h,t}v$  we obtain that

$$(2.69) \quad \mathcal{Q}[\hat{v} - \widehat{\mathcal{R}_{h,t}v}] = \inf_{\varphi \in X^h} \mathcal{Q}[\hat{v} - \varphi].$$

Using property (1.21) we obtain that

$$(2.70) \quad \mathcal{Q}[\hat{v} - \widehat{\mathcal{R}_{h,t}v}] \leq \lambda_{\max}(Q) C_X^2 h^2 \|\hat{v}\|_{W_D^{2,2}(T)}^2.$$

It is easy to obtain that

- for  $n = 1$ ,

$$(2.71) \quad \lambda_{\max}(Q) \leq \|a \nabla F\|_{L^\infty(\Omega_T)};$$

- for  $n \geq 2$ ,

$$(2.72) \quad \lambda_{\max}(Q) \leq C_n \mu \sigma^{\frac{n-1}{2}} \|(\text{Trace}[\sigma])^{-1}\|_{L^\infty(\Omega_T)}^{\frac{n}{2}-1}. \quad \square$$

LEMMA 2.10. Assume that  $\partial_t a \equiv 0$ ,  $\Omega$  is convex,  $\beta_\sigma < 1$ , and  $(\text{Trace}[\sigma])^{-1} \in L^\infty(\Omega_T)$ , and then

$$(2.73) \quad \mathcal{A}_T[\rho] \leq C h^2 \|g\|_{L^2(\Omega_T)}^2.$$

Remark 2.9. The constant  $C$  depends on  $C_X$ ,  $n$ ,  $\Omega$ ,  $\lambda_{\min}(a)$ , and  $\|(\text{Trace}[\sigma])^{-1}\|_{L^\infty(\Omega_T)}$ . If  $n \geq 5$  it also depends on  $\|\text{Trace}[\sigma]\|_{L^\infty(\Omega_T)}$ , and if  $n = 1$  it also depends on  $\lambda_{\max}(a)$ .

Proof. The proof is a straightforward application of Lemma 2.9 and Theorem 1.1. Observe that in dimension one  $a \nabla F = (\int_\Omega a^{-1})^{-1}$ .  $\square$

LEMMA 2.11. Assume that  $\partial_t a \equiv 0$ ,  $\Omega$  is convex,  $\beta_\sigma < 1$ , and  $(\text{Trace}[\sigma])^{-1} \in L^\infty(\Omega_T)$ , and then

$$(2.74) \quad \|\rho\|_{L^2(\Omega_T)} \leq C h^2 \|g\|_{L^2(\Omega_T)}.$$

Remark 2.10. The constant  $C$  depends on  $C_X$ ,  $n$ ,  $\Omega$ ,  $\lambda_{\min}(a)$ , and  $\|(\text{Trace}[\sigma])^{-1}\|_{L^\infty(\Omega_T)}$ . If  $n \geq 5$  it also depends on  $\|\text{Trace}[\sigma]\|_{L^\infty(\Omega_T)}$ , and if  $n = 1$  it also depends on  $\lambda_{\max}(a)$ .

Proof. The proof follows from standard duality techniques (see, for instance, Theorem 5.7.6 of [18]). We choose  $v \in L^2(0, T, H_0^1(\Omega))$  to be the solution of the following linear problem: For all  $w \in L^2(0, T, H_0^1(\Omega))$

$$(2.75) \quad \mathcal{A}_T[w, v] = (w, \rho)_{L^2(\Omega_T)}.$$

Choosing  $w = \rho$  in (2.75) we deduce that

$$(2.76) \quad \|\rho\|_{L^2(\Omega_T)}^2 = \mathcal{A}_T[\rho, v - \mathcal{R}_{h,t}v].$$

Using the Cauchy–Schwartz inequality we deduce that

$$(2.77) \quad \|\rho\|_{L^2(\Omega_T)}^2 \leq (\mathcal{A}_T[\rho])^{\frac{1}{2}} (\mathcal{A}_T[v - \mathcal{R}_{h,t}v])^{\frac{1}{2}}.$$

Using Theorem 1.1 we obtain that

$$(2.78) \quad \|\hat{v}\|_{L^2(0,T,W_D^{2,2}(\Omega))} \leq C \|\rho\|_{L^2(\Omega_T)}.$$

Using Lemma 2.9 we obtain that

$$(2.79) \quad (\mathcal{A}_T[v - \mathcal{R}_{h,t}v])^{\frac{1}{2}} \leq Ch\|\rho\|_{L^2(\Omega_T)}.$$

It follows that

$$(2.80) \quad \|\rho\|_{L^2(\Omega_T)} \leq Ch(\mathcal{A}_T[\rho])^{\frac{1}{2}}.$$

We deduce the lemma by applying Lemma 2.10 to bound  $\mathcal{A}_T[\rho]$ .  $\square$

The proof of Theorem 1.3 is a straightforward application of Lemmas 2.2, 2.6, 2.8, 2.10, and 2.11.

### 2.2.2. Medium with a continuum of time scales: Proof of Theorem 1.4.

In this subsection we will assume that the finite elements are in  $H^2(\Omega) \cap H_0^1(\Omega)$  and satisfy inverse inequality (1.22).

LEMMA 2.12.

$$(2.81) \quad \frac{1}{2} \|(u - u_h)(t)\|_{L^2(\Omega)}^2 + \mathcal{A}_t[u - u_h] = \int_{\Omega_t} \frac{\hat{\rho}}{|\det \nabla F| \circ F^{-1}} \left( \hat{g} + \sum_{i,j=1}^n \sigma_{i,j} \circ F^{-1} \partial_i \partial_j \hat{u}_h - \partial_t \hat{u}_h \right).$$

*Proof.* Consider (2.55). We have

$$(2.82) \quad \int_{\Omega_t} \rho \partial_t (u - u_h) = \int_{\Omega_t} \frac{\hat{\rho}}{|\det \nabla F| \circ F^{-1}} \partial_t (\hat{u} - \hat{u}_h) + \int_{\Omega_t} \rho \partial_t F(\nabla F)^{-1} \nabla (u - u_h).$$

Using (1.2) we obtain that

$$(2.83) \quad \int_{\Omega_t} \rho \partial_t F(\nabla F)^{-1} \nabla (u - u_h) = -\mathcal{A}_t[\rho, u - u_h] - \sum_{i,j=1}^n \int_{\Omega_t} \hat{\rho} Q_{i,j} \partial_i \partial_j (\hat{u} - \hat{u}_h). \quad \square$$

LEMMA 2.13.

$$(2.84) \quad \left\| \frac{\partial_t \hat{u}_h}{|\det(\nabla F)|^{\frac{1}{2}} \circ F^{-1}} \right\|_{L^2(\Omega_T)} \leq 2\|g\|_{L^2(\Omega_T)} + C\|\hat{u}_h\|_{L^2(0,T,W_D^{2,2}(\Omega))},$$

where the constant  $C$  depends on  $n$ ,  $\lambda_{\max}(a)$ ,  $\|\text{Trace}[\sigma]\|_{L^\infty(\Omega_T)}$ , and  $\mu_\sigma$ .

*Proof.* Using the change of variable  $y = F(x, t)$  in (1.27) we obtain that for all  $\varphi \in X^h$

$$(2.85) \quad \begin{cases} (\varphi, \frac{\partial_t \hat{u}_h}{|\det(\nabla F)| \circ F^{-1}})_{L^2(\Omega)} = \sum_{i,j=1}^n \int_{\Omega} (\varphi, Q_{i,j} \partial_i \partial_j \hat{u}_h)_{L^2(\Omega)} \\ \quad + (\varphi, \frac{\hat{g}}{|\det(\nabla F)| \circ F^{-1}})_{L^2(\Omega)}, \\ \hat{u}_h(x, 0) = 0. \end{cases}$$

Recall that  $Q$  is given by (2.68). We choose  $\varphi = \partial_t \hat{u}$  and observe that

$$(2.86) \quad \frac{\sigma}{|\det \nabla F|^{\frac{1}{2}}} = \frac{\sigma}{|\det \sigma|^{\frac{1}{4}}} |\det a|^{\frac{1}{4}}.$$

Thus

$$(2.87) \quad \left\| \frac{\sigma}{|\det \nabla F|^{\frac{1}{2}}} \right\| \leq C(n, \lambda_{\max}(a), \|\text{Trace}[\sigma]\|_{L^\infty(\Omega_T)}, \mu_\sigma).$$

We deduce the lemma by the Minkowski inequality.  $\square$

Combining Lemmas 2.12 and 2.13 we obtain the following lemma.

LEMMA 2.14.

$$(2.88) \quad \frac{1}{2} \|(u - u_h)(T)\|_{L^2(\Omega)}^2 + \mathcal{A}_T[u - u_h] \leq \|\rho\|_{L^2(\Omega_T)} \left( \|g\|_{L^2(\Omega_T)} + C \|\hat{u}_h\|_{L^2(0,T,W_D^{2,2}(\Omega))} \right),$$

where the constant  $C$  depends on  $n$ ,  $\lambda_{\max}(a)$ ,  $\|\text{Trace}[\sigma]\|_{L^\infty(\Omega_T)}$ , and  $\mu_\sigma$ .

LEMMA 2.15. Assume that  $\Omega$  is convex and Condition 1.1 is satisfied; then

$$(2.89) \quad \|\rho\|_{L^2(\Omega_T)} \leq Ch^2 \|g\|_{L^2(\Omega_T)}.$$

Remark 2.11. The constant  $C$  depends on  $C_X$ ,  $n$ ,  $\Omega$ ,  $\delta$ ,  $\epsilon$ ,  $\lambda_{\min}(a)$ , and  $\lambda_{\max}(a)$ .

Proof. The proof is similar to the proof of Lemma 2.11. As in the proof of Lemma 2.11 we choose  $v \in L^2(0, T, H_0^1(\Omega))$  to be the solution of the following linear problem: For all  $w \in L^2(0, T, H_0^1(\Omega))$

$$(2.90) \quad A_T[w, v] = (w, \rho)_{L^2(\Omega_T)}.$$

Choosing  $w = \rho$  in (2.90) we deduce that

$$(2.91) \quad \|\rho\|_{L^2(\Omega_T)}^2 = \mathcal{A}_T[\rho, v - \mathcal{R}_{h,t}v].$$

Using the Cauchy–Schwartz inequality we deduce that

$$(2.92) \quad \|\rho\|_{L^2(\Omega_T)}^2 \leq (\mathcal{A}_T[\rho])^{\frac{1}{2}} (\mathcal{A}_T[v - \mathcal{R}_{h,t}v])^{\frac{1}{2}}.$$

Using Theorem 1.2 we obtain that

$$(2.93) \quad \|\hat{v}\|_{L^2(0,T,W_D^{2,2}(\Omega))} \leq C \|\rho\|_{L^2(\Omega_T)}.$$

Using Lemma 2.9 we obtain that

$$(2.94) \quad (\mathcal{A}_T[v - \mathcal{R}_{h,t}v])^{\frac{1}{2}} \leq Ch \|\rho\|_{L^2(\Omega_T)}.$$

It follows that

$$(2.95) \quad \|\rho\|_{L^2(\Omega)} \leq Ch (\mathcal{A}_T[\rho])^{\frac{1}{2}}.$$

We deduce the lemma by applying Lemma 2.9 and Theorem 1.2 to bound  $A_T[\rho]$ .  $\square$

LEMMA 2.16. Assume that  $\Omega$  is convex and that  $\text{Trace}[\sigma] \in L^\infty(\Omega_T)$ .

$$(2.96) \quad \|\hat{u}_h\|_{L^2(0,T,W_D^{2,2}(\Omega))} \leq \frac{C}{h} \|g\|_{L^2(\Omega_T)}.$$

Remark 2.12. The constant  $C$  depends on  $C_X$ ,  $n$ ,  $\Omega$ ,  $\lambda_{\min}(a)$ ,  $\lambda_{\max}(a)$ , and  $\|\text{Trace}[\sigma]\|_{L^\infty(\Omega_T)}$ .

Proof. Using the inverse inequality (1.22) of the finite elements we obtain that

$$(2.97) \quad \|\hat{u}_h\|_{L^2(0,T,W_D^{2,2}(\Omega))} \leq \frac{C_X}{h} \|\nabla \hat{u}_h\|_{L^2(0,T,W_D^{2,2}(\Omega))}.$$

Using the change of variables  $y = F(x)$  we obtain that

$$(2.98) \quad \|\nabla \hat{u}_h\|_{L^2(0,T,W_D^{2,2}(\Omega))}^2 \leq C \mathcal{A}_T[u_h],$$

where  $C$  depends on  $n$ ,  $\lambda_{\min}(a)$ , and  $\|\text{Trace}[\sigma]\|_{L^\infty(\Omega_T)}$ . We deduce the lemma by using Lemma 2.7.  $\square$



The proof of Theorem 1.4 is a straightforward application of Lemmas 2.14, 2.15, and 2.16.

**2.2.3. Numerical homogenization in time and space: Proof of Theorems 1.5–1.7.** We use the notation of subsection 1.3. First we will prove Theorem 1.5; let us observe that the numerical scheme associated to (1.34) is stable. Indeed choosing  $\psi = v_{n+1}$  one gets

$$(2.99) \quad \begin{aligned} |v_{n+1}(t_{n+1})|_{L^2(\Omega)}^2 &= (v_{n+1}(t_n), v_n(t_n))_{L^2(\Omega)} \\ &+ \frac{1}{2}(|v_{n+1}(t_{n+1})|_{L^2(\Omega)}^2 - |v_{n+1}(t_n)|_{L^2(\Omega)}^2) \\ &- \int_{t_n}^{t_{n+1}} \left( a[v_{n+1}(t)] + (v_{n+1}(t), g(t))_{L^2(\Omega)} \right) dt. \end{aligned}$$

It follows by the Cauchy–Schwartz inequality that

$$(2.100) \quad \begin{aligned} \frac{1}{2}|v_{n+1}(t_{n+1})|_{L^2(\Omega)}^2 &\leq \frac{1}{2}|v_n(t_n)|_{L^2(\Omega)}^2 \\ &- \int_{t_n}^{t_{n+1}} \left( a[v_{n+1}(t)] + (v_{n+1}(t), g(t))_{L^2(\Omega)} \right) dt. \end{aligned}$$

Hence using Poincaré and Minkowski inequalities one obtains that

$$(2.101) \quad \begin{aligned} |v_{n+1}(t_{n+1})|_{L^2(\Omega)}^2 &+ \int_{t_n}^{t_{n+1}} a[v_{n+1}(t)] dt \leq |v_n(t_n)|_{L^2(\Omega)}^2 \\ &+ \frac{C_{n,\Omega}}{\lambda_{\min}(a)} \int_{t_n}^{t_{n+1}} |g(t)|_{L^2(\Omega)}^2 dt, \end{aligned}$$

which implies Theorem 1.5 and the stability of the scheme.

Integrating (1.27) with respect to time we obtain that, for  $\psi \in U_T^h$ ,

$$(2.102) \quad \begin{aligned} (\psi(t_{n+1}), u_h(t_{n+1}))_{L^2(\Omega)} &= (\psi(t_n), u_h(t_n))_{L^2(\Omega)} \\ &+ \int_{t_n}^{t_{n+1}} \left( (\partial_t \psi(t), u_h(t))_{L^2(\Omega)} \right. \\ &\quad \left. - a[\psi(t), u_h(t)] + (\psi(t), g(t))_{L^2(\Omega)} \right) dt. \end{aligned}$$

Let us write  $(u^i)$  the coordinates of  $u_h$  associated to the basis  $(\varphi_i \circ F)$ , i.e.,

$$(2.103) \quad u_h(x, t) := \sum_i u^i(t) \varphi_i(F(x, t)),$$

and define

$$(2.104) \quad u_n(x, t) := \sum_i u^i(t_n) \varphi_i(F(x, t)).$$

Subtracting (2.102) and (1.34) we obtain that, for  $\psi \in Z_T$ ,

$$(2.105) \quad \begin{aligned} (\psi(t_{n+1}), (u_{n+1} - v_{n+1})(t_{n+1}))_{L^2(\Omega)} &= (\psi(t_n), (u_n - v_n)(t_n))_{L^2(\Omega)} \\ &+ \int_{t_n}^{t_{n+1}} \left( (\partial_t \psi(t), (u_h - v_{n+1})(t))_{L^2(\Omega)} \right. \\ &\quad \left. - a[\psi(t), (u_h - v_{n+1})(t)] \right) dt. \end{aligned}$$

Choosing  $\psi = u_{n+1} - v_{n+1}$  we deduce using the Cauchy–Schwartz inequality that

$$\begin{aligned}
 (2.106) \quad & \frac{1}{2} |(u_{n+1} - v_{n+1})(t_{n+1})|_{L^2(\Omega)}^2 + \int_{t_n}^{t_{n+1}} a[(u_{n+1} - v_{n+1})(t)] dt \\
 & \leq \frac{1}{2} |(u_n - v_n)(t_n)|_{L^2(\Omega)}^2 + \int_{t_n}^{t_{n+1}} \left( (\partial_t(u_{n+1} - v_{n+1})(t), (u_h - u_{n+1})(t))_{L^2(\Omega)} \right. \\
 & \quad \left. - a[(u_{n+1} - v_{n+1})(t), (u_h - u_{n+1})(t)] \right) dt.
 \end{aligned}$$

*Time-independent medium.* Observe that if the medium is time-independent, then (2.106) can be written

$$\begin{aligned}
 (2.107) \quad & \frac{1}{2} |(u_{n+1} - v_{n+1})(t_{n+1})|_{L^2(\Omega)}^2 + \int_{t_n}^{t_{n+1}} a[(u_{n+1} - v_{n+1})(t)] dt \\
 & \leq \frac{1}{2} |(u_n - v_n)(t_n)|_{L^2(\Omega)}^2 - \int_{t_n}^{t_{n+1}} a[(u_{n+1} - v_{n+1})(t), (u_h - u_{n+1})(t)] dt,
 \end{aligned}$$

which leads us to

$$\begin{aligned}
 (2.108) \quad & \frac{1}{2} |(u_{n+1} - v_{n+1})(t_{n+1})|_{L^2(\Omega)}^2 + \int_{t_n}^{t_{n+1}} a[(u_{n+1} - v_{n+1})(t)] dt \\
 & \leq \frac{1}{2} |(u_n - v_n)(t_n)|_{L^2(\Omega)}^2 \\
 & \quad + \int_{t_n}^{t_{n+1}} \int_{t_n}^{t_{n+1}} 1(t < s) a[(u_{n+1} - v_{n+1})(t), \partial_s u_h(s)] ds dt.
 \end{aligned}$$

Write  $\Delta t := t_{n+1} - t_n$ . Using the Minkowski inequality we obtain that

$$\begin{aligned}
 (2.109) \quad & a[(u_{n+1} - v_{n+1})(t), \partial_s u_h(s)] \leq \frac{1}{2\Delta t} a[(u_{n+1} - v_{n+1})(t)] \\
 & \quad + \frac{1}{2} \Delta t a[\partial_s u_h(s)].
 \end{aligned}$$

It follows from (2.108) that

$$\begin{aligned}
 (2.110) \quad & |(u_{n+1} - v_{n+1})(t_{n+1})|_{L^2(\Omega)}^2 + \int_{t_n}^{t_{n+1}} a[(u_{n+1} - v_{n+1})(t)] dt \\
 & \leq |(u_n - v_n)(t_n)|_{L^2(\Omega)}^2 + |\Delta t|^2 \int_{t_n}^{t_{n+1}} a[\partial_s u_h(s)] ds.
 \end{aligned}$$

Observing that

$$\begin{aligned}
 (2.111) \quad & \int_{t_n}^{t_{n+1}} a[(u_{n+1} - v_{n+1})(t)] dt \geq 0.5 \int_{t_n}^{t_{n+1}} a[(u_h - v_{n+1})(t)] dt \\
 & \quad - \int_{t_n}^{t_{n+1}} a[(u_h - u_{n+1})(t)] dt
 \end{aligned}$$

and

$$(2.112) \quad \int_{t_n}^{t_{n+1}} a[(u_h - u_{n+1})(t)] dt \leq |\Delta t|^2 \int_{t_n}^{t_{n+1}} a[\partial_s u_h(s)] ds,$$

we obtain that

$$(2.113) \quad \begin{aligned} & |(u_{n+1} - v_{n+1})(t_{n+1})|_{L^2(\Omega)}^2 + 0.5 \int_{t_n}^{t_{n+1}} a[(u_h - v_{n+1})(t)] dt \\ & \leq |(u_n - v_n)(t_n)|_{L^2(\Omega)}^2 + \frac{3}{2} |\Delta t|^2 \int_{t_n}^{t_{n+1}} a[\partial_s u_h(s)] ds. \end{aligned}$$

In conclusion we have obtained the following lemma.

LEMMA 2.17. *Let  $v \in Z_T$  be the solution of (1.34). We have*

$$(2.114) \quad \|(u_h - v)(T)\|_{L^2(\Omega)}^2 + \int_0^T a[(u_h - v)(t)] dt \leq 3|\Delta t|^2 \int_0^T a[\partial_s u_h(s)] ds.$$

Combining Lemma 2.3 with Lemma 2.17 we obtain Theorem 1.6.

*Time-dependent medium.* Observe that

$$\partial_t(u_{n+1} - v_{n+1}) = \partial_t F(\nabla F)^{-1} \nabla(u_{n+1} - v_{n+1}).$$

It follows after writing  $\partial_t F = \operatorname{div} a \nabla F$ , integration by parts, and using the change of variables  $y = F(x, t)$  in (2.106) that

$$(2.115) \quad \begin{aligned} & \frac{1}{2} |(u_{n+1} - v_{n+1})(t_{n+1})|_{L^2(\Omega)}^2 + \int_{t_n}^{t_{n+1}} a[(u_{n+1} - v_{n+1})(t)] dt \\ & \leq \frac{1}{2} |(u_n - v_n)(t_n)|_{L^2(\Omega)}^2 - 2 \int_{t_n}^{t_{n+1}} a[(u_{n+1} - v_{n+1})(t), (u_h - u_{n+1})(t)] dt \\ & \quad - \sum_{i,j} \int_{t_n}^{t_{n+1}} \int_{\Omega} (\hat{u}_h - \hat{u}_{n+1}) Q_{i,j} \partial_i \partial_j (\hat{u}_{n+1} - \hat{v}_{n+1}) dt dy. \end{aligned}$$

Hence using the Minkowski inequality we obtain that

$$(2.116) \quad \begin{aligned} & |(u_{n+1} - v_{n+1})(t_{n+1})|_{L^2(\Omega)}^2 + \int_{t_n}^{t_{n+1}} a[(u_{n+1} - v_{n+1})(t)] dt \\ & \leq |(u_n - v_n)(t_n)|_{L^2(\Omega)}^2 + 4 \int_{t_n}^{t_{n+1}} a[(u_h - u_{n+1})(t)] dt \\ & \quad - 2 \sum_{i,j} \int_{t_n}^{t_{n+1}} \int_{\Omega} (\hat{u}_h - \hat{u}_{n+1}) Q_{i,j} \partial_i \partial_j (\hat{u}_{n+1} - \hat{v}_{n+1}) dt dy. \end{aligned}$$

Using the Minkowski inequality we obtain that

$$\begin{aligned}
 (2.117) \quad & \left| \sum_{i,j} \int_{t_n}^{t_{n+1}} \int_{\Omega} (\hat{u}_h - \hat{u}_{n+1}) Q_{i,j} \partial_i \partial_j (\hat{u}_{n+1} - \hat{v}_{n+1}) dt dy \right| \\
 & \leq C_A n^2 \int_{t_n}^{t_{n+1}} \int_{\Omega} |\hat{u}_h - \hat{u}_{n+1}|^2 dt dy \\
 & \quad + \frac{\lambda_{\max}(Q)}{C_A} \int_{t_n}^{t_{n+1}} \sum_{i,j} \int_{\Omega} |\partial_i \partial_j (\hat{u}_{n+1} - \hat{v}_{n+1})|^2 dt dy.
 \end{aligned}$$

Using the inverse inequality (1.22) and the change of variable  $y = F(x)$  we obtain that

$$\begin{aligned}
 (2.118) \quad & \int_{t_n}^{t_{n+1}} \sum_{i,j} \int_{\Omega} |\partial_i \partial_j (\hat{u}_{n+1} - \hat{v}_{n+1})|^2 dt dy \leq \frac{C_X}{h^2 \lambda_{\min}(Q)} \\
 & \int_{t_n}^{t_{n+1}} a[(u_{n+1} - v_{n+1})(t)] dt.
 \end{aligned}$$

In resume, choosing  $C_A = \frac{4C_X \lambda_{\max}(Q)}{h^2 \lambda_{\min}(Q)}$  we have obtained that

$$\begin{aligned}
 (2.119) \quad & |(u_{n+1} - v_{n+1})(t_{n+1})|_{L^2(\Omega)}^2 + 0.5 \int_{t_n}^{t_{n+1}} a[(u_{n+1} - v_{n+1})(t)] dt \\
 & \leq |(u_n - v_n)(t_n)|_{L^2(\Omega)}^2 + 8 \int_{t_n}^{t_{n+1}} a[(u_h - u_{n+1})(t)] dt \\
 & \quad + \frac{8C_X \lambda_{\max}(Q)}{h^2 \lambda_{\min}(Q)} n^2 \int_{t_n}^{t_{n+1}} \int_{\Omega} |\hat{u}_h - \hat{u}_{n+1}|^2 dt dy.
 \end{aligned}$$

A computation similar to the one leading to (2.109) gives us

$$\begin{aligned}
 (2.120) \quad & |(u_{n+1} - v_{n+1})(t_{n+1})|_{L^2(\Omega)}^2 + \frac{1}{4} \int_{t_n}^{t_{n+1}} a[(u_h - v_{n+1})(t)] dt \\
 & \leq |(u_n - v_n)(t_n)|_{L^2(\Omega)}^2 + 9 \int_{t_n}^{t_{n+1}} a[(u_h - u_{n+1})(t)] dt \\
 & \quad + \frac{8C_X \lambda_{\max}(Q)}{h^2 \lambda_{\min}(Q)} n^2 \int_{t_n}^{t_{n+1}} \int_{\Omega} |\hat{u}_h - \hat{u}_{n+1}|^2 dt dy.
 \end{aligned}$$

Moreover using the change of variables  $F(x) = y$  and the inverse inequality (1.23) we obtain that

$$(2.121) \quad \int_{t_n}^{t_{n+1}} a[(u_h - u_{n+1})(t)] dt \leq \frac{C_X \lambda_{\max}(Q)}{h^2} \int_{t_n}^{t_{n+1}} \int_{\Omega} |\hat{u}_h - \hat{u}_{n+1}|^2 dt dy.$$

Let us also observe that

$$(2.122) \quad \int_{t_n}^{t_{n+1}} \int_{\Omega} |\hat{u}_h - \hat{u}_{n+1}|^2 dt dy \leq |\Delta t|^2 \int_{t_n}^{t_{n+1}} \int_{\Omega} |\partial_t \hat{u}_h|^2 dt dy.$$

It follows that

$$(2.123) \quad \begin{aligned} & |(u_{n+1} - v_{n+1})(t_{n+1})|_{L^2(\Omega)}^2 + \frac{1}{4} \int_{t_n}^{t_{n+1}} a[(u_h - v_{n+1})(t)] dt \\ & \leq |(u_n - v_n)(t_n)|_{L^2(\Omega)}^2 + C_B \frac{|\Delta t|^2}{h^2} \int_{t_n}^{t_{n+1}} \int_{\Omega} |\partial_t \hat{u}_h|^2 dt dy, \end{aligned}$$

with

$$(2.124) \quad C_B = C_n C_X \lambda_{\max}(Q) \left( 1 + \frac{1}{\lambda_{\min}(Q)} \right).$$

We deduce that

$$(2.125) \quad \|(u_h - v)(T)\|_{L^2(\Omega)}^2 + \frac{1}{4} \int_0^T a[(u_h - v)(t)] dt \leq C_B \frac{|\Delta t|^2}{h^2} \int_{\Omega_T} |\partial_t \hat{u}_h|^2 dt dy.$$

Using Lemma 2.4 to control  $C_B$  and combining (2.125) with Theorem 1.2 we obtain Theorem 1.7.

**3. Numerical experiments.** The purpose of this section is to give several illustrations of the implementation and performance of the numerical method. The computational domain is the unit square in dimension two. Equation (1.1) is solved on a fine tessellation characterized by 16129 interior nodes (degrees of freedom).

Three different coarse tessellations with 9, 49, and 225 degrees of freedoms are considered.

The parabolic operator associated to (1.1) has been homogenized onto these coarse meshes using the method presented in subsection 1.3. We have chosen splines to be the space  $X^h$  in subsection 1.2.

### 3.1. Time-independent examples.

*Example 3.1* (time-independent site percolation). In this example we consider the site-percolating medium.  $a$  is set to be equal to 1 or 100 with probability 1/2 on each triangle of a fine mesh. (1.1) has been solved with  $g = 1$  and  $g = \sin(2.4x - 1.8y + 2\pi t)$ . Figure 3.1 shows  $u$  computed on 16641 interior nodes and  $u_h$  computed on 9 interior nodes in the case  $g = 1$  at time 1. The fine mesh and coarse mesh errors are given in Table 3.1.

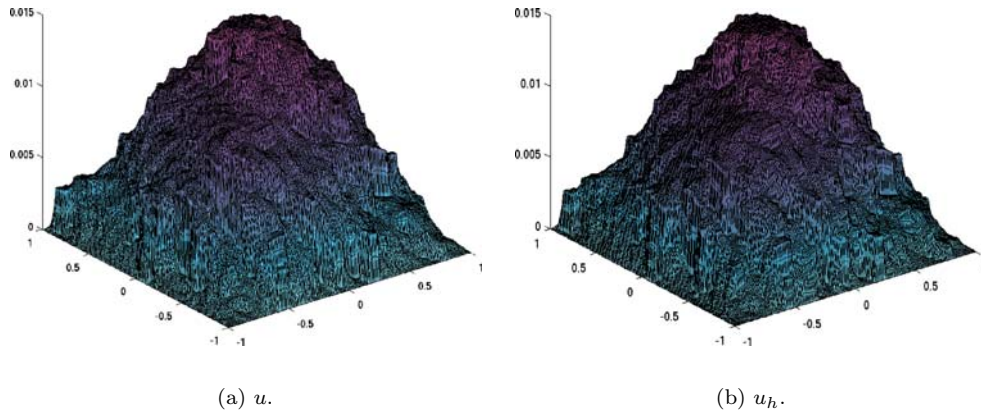


FIG. 3.1.  $u$  computed on 16641 interior nodes and  $u_h$  computed on 9 interior nodes.

TABLE 3.1

Experiment	dof	$L^1$	$L^\infty$	$L^2$	$H^1$
Coarse mesh error. Time-independent percolation with $g = 1$ .	9	0.0142	0.0389	0.0168	0.0366
	49	0.0077	0.0450	0.0101	0.0482
	225	0.0035	0.0228	0.0060	0.0293
Fine mesh error. Time-independent percolation with $g = 1$ .	9	0.0196	0.0843	0.0251	0.1193
	49	0.0136	0.0698	0.0184	0.1028
	225	0.0040	0.0243	0.0070	0.0485
Coarse mesh error. Time-independent percolation with $g = \sin(2.4x - 1.8y + 2\pi t)$ .	9	0.0236	0.0569	0.0262	0.0477
	49	0.0181	0.0571	0.0215	0.0558
	225	0.0119	0.0774	0.0167	0.0939
Fine mesh error. Time-independent percolation with $g = \sin(2.4x - 1.8y + 2\pi t)$ .	9	0.0424	0.1099	0.0512	0.1712
	49	0.0277	0.0985	0.0348	0.1451
	225	0.0174	0.0886	0.0242	0.1192
Coarse mesh error. Multiscale trigonometric time-dependent. $g = 1$ .	9	0.0018	0.0045	0.0019	0.0039
	49	0.0012	0.0054	0.0015	0.0060
Fine mesh error. Multiscale trigonometric time-dependent. $g = 1$ .	9	0.0031	0.0096	0.0034	0.0242
	49	0.0014	0.0059	0.0016	0.0166
Coarse mesh error. Multiscale trigonometric time-dependent. $g = \sin(2.4x - 1.8y + 2\pi t)$ .	9	0.0043	0.0087	0.0044	0.0085
	49	0.0033	0.0079	0.0035	0.0084
Fine mesh error. Multiscale trigonometric time-dependent medium. $g = \sin(2.4x - 1.8y + 2\pi t)$ .	9	0.0082	0.0199	0.0087	0.0379
	49	0.0038	0.0104	0.0040	0.0244
Coarse mesh error. Time-dependent random fractal.	9	0.0046	0.0074	0.0052	0.0065
	49	0.0036	0.0046	0.0036	0.0059
Fine mesh error. Time-dependent random fractal.	9	0.0039	0.0082	0.0043	0.0222
	49	0.0033	0.0054	0.0034	0.0168

**3.2. Time-dependent examples.** In the following examples we consider media characterized by a continuum of time scales.

*Example 3.2* (Time-dependent multiscale trigonometric). In this example  $a$  is given by (1.20). Although the number of fine time steps to solve (1.1) is 2663, only 134 coarse time steps have been used to solve the homogenized equation. Hence if one also takes into account homogenization in space, the compression factor is of the order of 35000 for the coarse mesh with 9 interior nodes.

Figure 3.2 shows the curves of  $t \rightarrow a(x_0, t)$  and  $t \rightarrow F(x_0, t)$  for a given  $x_0 \in \Omega$  and the coarse and fine mesh relative to  $L^2$  and  $H^1$  errors with respect to time. The initial increase of the relative error is due to the initial value  $u(x, 0) \equiv 0$ .

The coarse and fine meshes errors at  $t = 0.1$  are given in Table 3.1 for  $g = 1$ . The errors at  $t = 0.1$  are given for  $g = \sin(2.4x - 1.8y + 2\pi t)$ .

*Example 3.3* (Time-dependent random fractal). In this case,  $a$  is given by a product of discontinuous functions oscillating randomly at different scales:  $a(x, t) = a_1(x, t)a_2(x, t) \dots a_n(x, t)$  and  $a_i(x, t) = c_{pq}$  for  $x \in [\frac{p}{2^i}, \frac{p+1}{2^i}) \times [\frac{q}{2^i}, \frac{q+1}{2^i})$  in the time interval  $0.1 \times [\frac{k}{4^i}, \frac{k+1}{4^i})$ .  $c_{pq}$  is uniformly random in  $[\frac{1}{\gamma}, \gamma]$ ,  $n = 6$ , and  $\gamma = 0.7$ . In this example, we have  $\frac{\lambda_{\max}(a)}{\lambda_{\min}(a)} = 160.3295$ . The number of fine time steps is 3482, and the number of coarse time steps is 175.

$a$  and the map  $(F_1, F_2)$  are drawn in Figure 3.3.  $L^2$  and  $H^1$  errors are given in Figure 3.4. Coarse and fine mesh errors at  $t = 0.1$  are given in Table 3.1. We have chosen  $g = 1$  in this numerical experiment, and one obtains similar results by choosing  $g = \sin(2.4x - 1.8y + 2\pi t)$ .

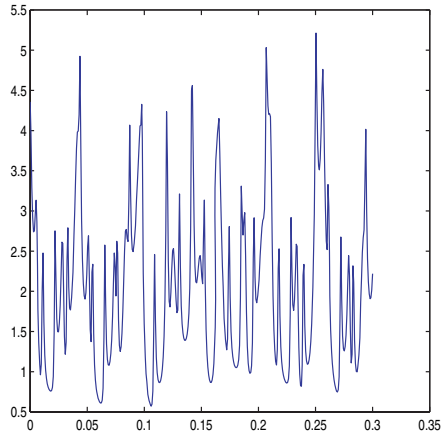
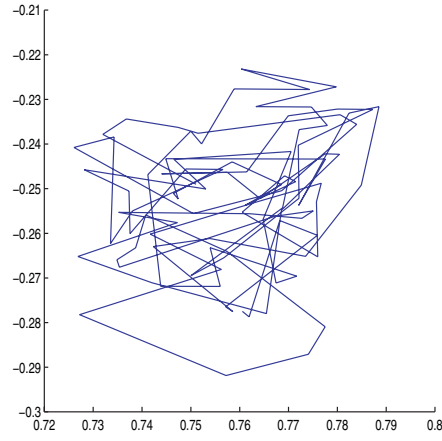
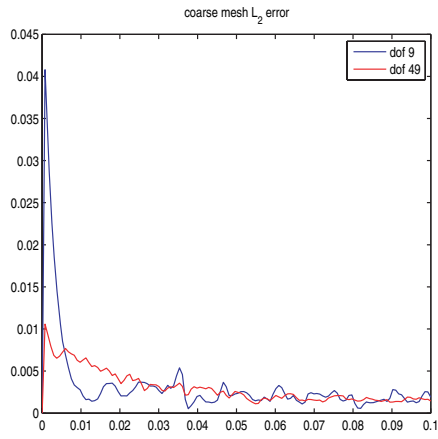
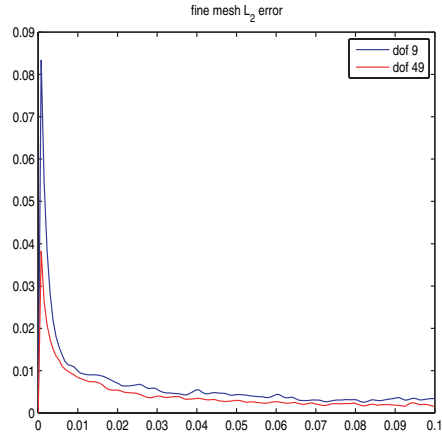
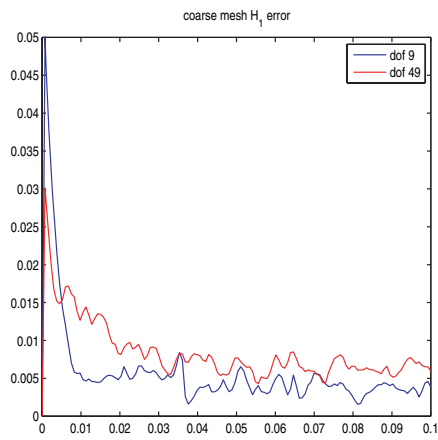
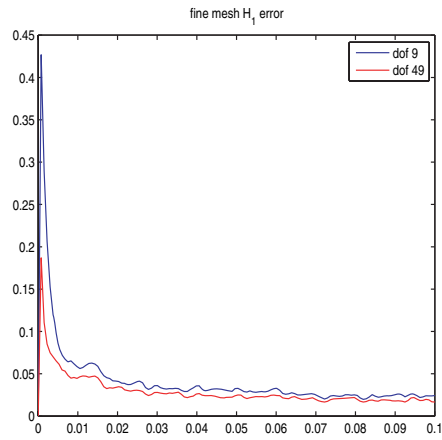
(a)  $t \rightarrow a(x_0, t)$ .(b) Top view of  $t \rightarrow F(x_0, t)$ .(c) Coarse mesh  $L^2$  error.(d) Fine Mesh  $L^2$  error.(e) Coarse Mesh  $H^1$  error.(f) Fine Mesh  $H^1$  error.

FIG. 3.2. Multiscale time-dependent trigonometric medium.

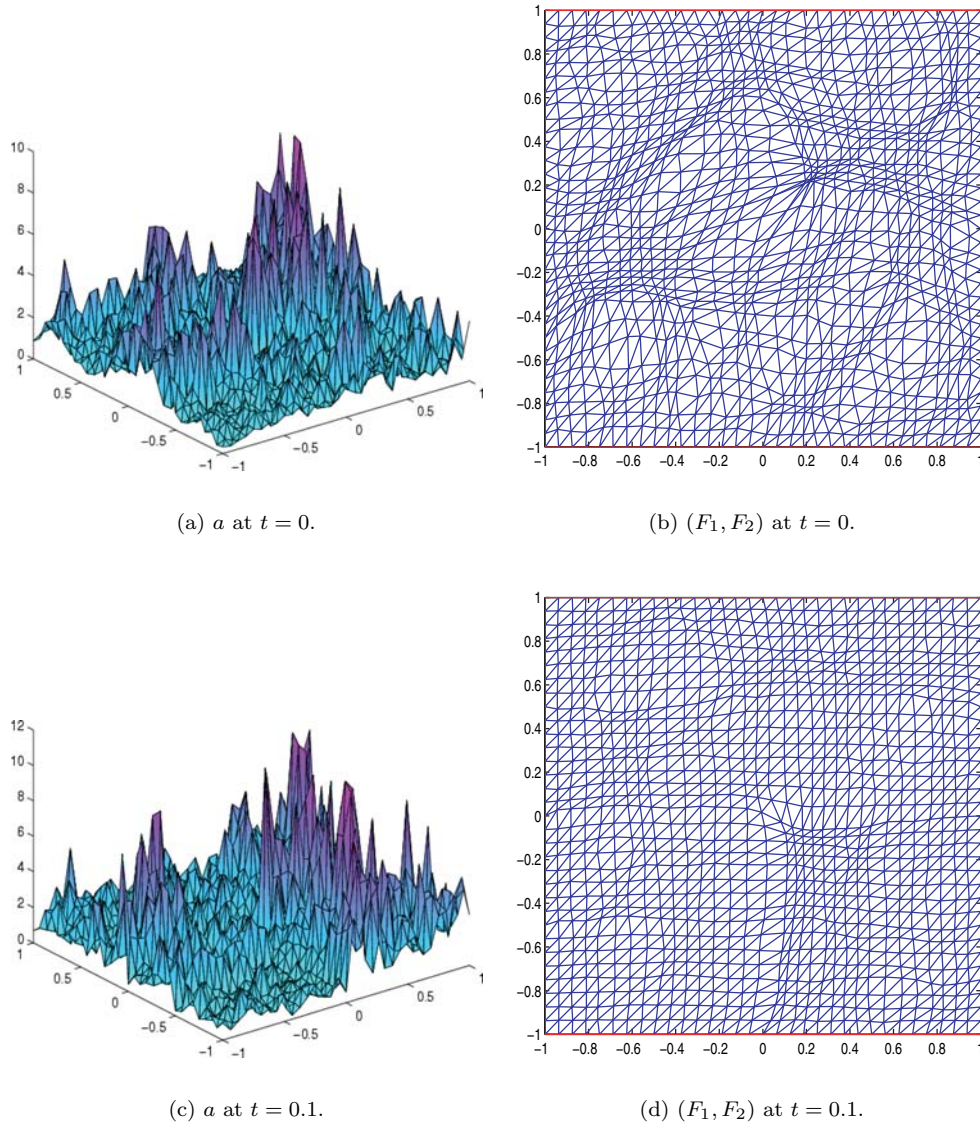
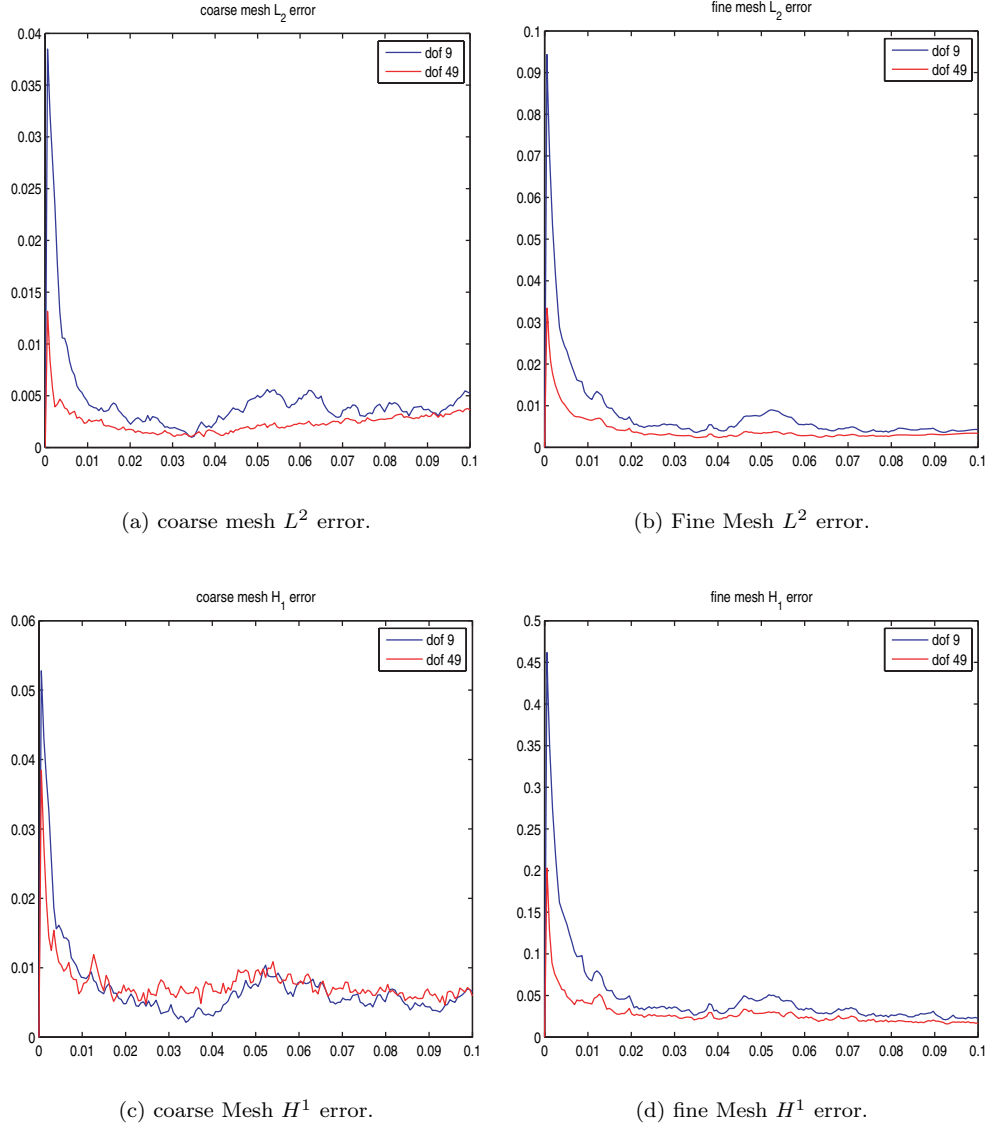


FIG. 3.3.  $a$  and  $(F_1, F_2)$  at time  $t = 0, t = 0.1$  for the time-dependent random fractal medium.



FIG. 3.4. Time-dependent random fractal medium at  $t = 0.1$ .

**Acknowledgment.** We thank two anonymous referees for detail comments and suggestions.

## REFERENCES

- [1] J. E. AARNES, *On the use of a mixed multiscale finite element method for greater flexibility and increased speed or improved accuracy in reservoir simulation*, Multiscale Model. Simul., 2 (2004), pp. 421–439.
- [2] G. ALBERTI AND S. MÜLLER, *A new approach to variational problems with multiple scales*, Comm. Pure Appl. Math., 54 (2001), pp. 761–825.
- [3] G. ALESSANDRINI AND V. NESI, *Univalent  $\sigma$ -harmonic mappings: Connections with quasiconformal mappings*, J. Anal. Math., 90 (2003), pp. 197–215.
- [4] G. ALLAIRE, *Shape Optimization by the Homogenization Method*, Springer, New York, 2001.
- [5] G. ALLAIRE AND R. BRIZZI, *A multiscale finite element method for numerical homogenization*, Multiscale Model. Simul., 4 (2005), pp. 790–812.
- [6] A. ANCONA, *Some results and examples about the behavior of harmonic functions and Green's functions with respect to second order elliptic operators*, Nagoya Math. J., 165 (2002), pp. 123–158.
- [7] A. AVERBUCH, G. BEYLKIN, R. COIFMAN, AND M. ISRAELI, *Multiscale inversion of elliptic operators*, in Signal and Image Representation in Combined Spaces, Wavelet Anal. Appl. 7, Academic Press, San Diego, CA, 1998, pp. 341–359.
- [8] I. BABUŠKA, G. CALOZ, AND J. E. OSBORN, *Special finite element methods for a class of second order elliptic problems with rough coefficients*, SIAM J. Numer. Anal., 31 (1994), pp. 945–981.
- [9] I. BABUŠKA AND J. E. OSBORN, *Generalized finite element methods: Their performance and their relation to mixed methods*, SIAM J. Numer. Anal., 20 (1983), pp. 510–536.
- [10] M. BEBENDORF, *Approximate inverse preconditioning of finite element discretizations of elliptic operators with nonsmooth coefficients*, SIAM J. Matrix. Anal., 27 (2006), pp. 909–929.
- [11] M. BEBENDORF, *Efficient inversion of Galerkin matrices of general second-order elliptic differential operators with nonsmooth coefficients*, Math. Comp., 74 (2005), pp. 1179–1199.
- [12] M. BEBENDORF, *Why Approximate Lu Decompositions of Finite Element Discretizations of Elliptic Operators can be Computed with Almost Linear Complexity*, preprint 8/2005, Max-Planck-Institute MiS, Leipzig, 2005.
- [13] M. BEBENDORF AND Y. CHEN, *Efficient Solution of Nonlinear Elliptic Problems Using Hierarchical Matrices with Broyden Updates*, preprint 51/2005, Max-Planck-Institut MiS, Leipzig, 2005.
- [14] M. BEBENDORF AND W. HACKBUSCH, *Existence of  $\mathcal{H}$ -matrix approximants to the inverse FE-matrix of elliptic operators with  $L^\infty$ -coefficients*, Numer. Math., 95 (2003), pp. 1–28.
- [15] G. BEN AROUS AND H. OWHADI, *Multiscale homogenization with bounded ratios and anomalous slow diffusion*, Comm. Pure Appl. Math., 56 (2003), pp. 80–113.
- [16] A. BENSOUSSAN, J. L. LIONS, AND G. PAPANICOLAOU, *Asymptotic Analysis for Periodic Structure*, North-Holland, Amsterdam, 1978.
- [17] G. BEYLKIN AND N. COULT, *A multiresolution strategy for reduction of elliptic PDEs and eigenvalue problems*, Appl. Comput. Harmon. Anal., 5 (1998), pp. 129–155.
- [18] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, 2nd ed., Springer, New York, 2002.
- [19] M. E. BREWSTER AND G. BEYLKIN, *A multiresolution strategy for numerical homogenization*, Appl. Comput. Harmon. Anal., 2 (1995), pp. 327–349.
- [20] F. BREZZI AND D. MARINI, *Subgrid phenomena and numerical schemes*, in Frontiers in Numerical Analysis (Durham, 2002), Universitext, Springer, Berlin, 2003, pp. 1–16.
- [21] M. BRIANE, *Nonlocal effects in two-dimension conductivity*, Arch. Ration. Mech. Anal., 182 (2006), pp. 255–267.
- [22] M. BRIANE, G. W. MILTON, AND V. NESI, *Change of sign of the corrector's determinant for homogenization in three-dimensional conductivity*, Arch. Ration. Mech. Anal., 173 (2004), pp. 133–150.
- [23] S. CAMPANATO, *Un risultato relativo ad equazioni ellittiche del secondo ordine di tipo non variazionale*, Ann. Scuola Norm. Sup. Pisa (3), 21 (1967), pp. 701–707.
- [24] S. CHEN, W. E, AND C.-W. SHU, *The heterogeneous multiscale method based on the discontinuous Galerkin method for hyperbolic and parabolic problems*, Multiscale Model. Simul., 3 (2005), pp. 871–894.
- [25] A. CHERTOCK AND D. LEVY, *On wavelet-based numerical homogenization*, Multiscale Model. Simul., 3 (2005), pp. 65–88.

- [26] R. COIFMAN, P.-L. LIONS, Y. MEYER, AND S. SEMMES, *Compensated compactness and Hardy spaces*, J. Math. Pures Appl. (9), 72 (1993), pp. 247–286.
- [27] M. DOROBANTU AND B. ENGQUIST, *Wavelet-based numerical homogenization*, SIAM J. Numer. Anal., 35 (1998), pp. 540–559.
- [28] W. E, B. ENGQUIST, X. LI, W. REN, AND E. VANDEN-EIJNDEN, *The Heterogeneous Multiscale Method: A Review*, preprint, Princeton, <http://www.math.princeton.edu/multiscale/review.pdf>.
- [29] Y. R. EFENDIEV, T. Y. HOU, AND X.-H. WU, *Convergence of a nonconforming multiscale finite element method*, SIAM J. Numer. Anal., 37 (2000), pp. 888–910.
- [30] A. ERN AND J.-L. GUERMOND, *Theory and Practice of Finite Element Methods*, Appl. Math. Sci. 159, Springer, New York, 2004. Two volumes bound as one.
- [31] L. EVANS, *Partial Differential Equations*, Grad. Stud. Math. 19, American Mathematical Society, Providence, RI, 1997.
- [32] C. FARHAT, I. HARARI, AND L. P. FRANCA, *The discontinuous enrichment method*, Comput. Methods Appl. Mech. Engrg., 190 (2001), pp. 6455–6479.
- [33] C. FARHAT, I. HARARI, AND U. HETMANIUK, *The discontinuous enrichment method for multiscale analysis*, Comput. Methods Appl. Mech. Engrg., 192 (2003), pp. 3195–3209.
- [34] V. N. FENCHENKO AND Ė. YA. KHRUSLOV, *Asymptotic behavior of solutions of differential equations with a strongly oscillating coefficient matrix that does not satisfy a uniform boundedness condition*, Dokl. Akad. Nauk Ukrain. SSR Ser. A, 95 (1981), pp. 24–27.
- [35] J. FISH AND A. WAGIMAN, *Multiscale finite element method for a locally nonperiodic heterogeneous medium*, Comput. Mech., 12 (1993), pp. 164–180.
- [36] J. FISH AND Y. ZHENG, *Multi-scale enrichment based on partition of unity*, Internat. J. Numer. Methods Engrg., (2005).
- [37] J. GARNIER, *Homogenization in a periodic and time-dependent potential*, SIAM J. Appl. Math., 57 (1997), pp. 95–111.
- [38] A. C. GILBERT, *A comparison of multiresolution and classical one-dimensional homogenization schemes*, Appl. Comput. Harmon. Anal., 5 (1998), pp. 1–35.
- [39] O. GOUBET, *Séparation des variables dans le problème de Stokes. Application à son approximation multiéchelles éléments finis*, C. R. Acad. Sci. Paris Sér. I Math., 315 (1992), pp. 1315–1318.
- [40] V. H. HOANG AND C. SCHWAB, *High-dimensional finite elements for elliptic problems with multiple scales*, Multiscale Model. Simul., 3 (2005), pp. 168–194.
- [41] T. Y. HOU AND X.-H. WU, *A multiscale finite element method for elliptic problems in composite materials and porous media*, J. Comput. Phys., 134 (1997), pp. 169–189.
- [42] T. Y. HOU, X.-H. WU, AND Y. ZHANG, *Removing the cell resonance error in the multiscale finite element method via a Petrov-Galerkin formulation*, Commun. Math. Sci., 2 (2004), pp. 185–205.
- [43] V. V. JIKOV, S. M. KOZLOV, AND O. A. OLEINIK, *Homogenization of Differential Operators and Integral Functionals*, Springer, Berlin, 1991.
- [44] Ė. YA. KHRUSLOV, *Homogenized models of composite media*, in Composite Media and Homogenization Theory (Trieste, 1990), Progr. Nonlinear Differential Equations Appl. 5, Birkhäuser Boston, Boston, MA, 1991, pp. 159–182.
- [45] C. KIPNIS AND S. R. S. VARADHAN, *Central limit theorem for additive functional of reversible markov processes and application to simple exclusion*, Comm. Math. Phys., 104 (1986), pp. 1–19.
- [46] S. M. KOZLOV, *The method of averaging and walks in inhomogeneous environments*, Russian Math. Surveys, 2 (1985), pp. 73–145.
- [47] S. M. KOZLOV, *Averaging of difference schemes*, Math. USSR. Sbornik, 2 (1987), pp. 351–369.
- [48] S. LARSON AND V. THOMEE, *Partial Differential Equations with Numerical Methods*, Texts Appl. Math. 45, Springer, New York, 2003.
- [49] S. H. LEE, P. JENNY, AND H. A. TCHELEPI, *A finite-volume method with hexahedral multiblock grids for modeling flow in porous media*, Comput. Geosci., 6 (2002), pp. 353–379.
- [50] S. H. LEE, P. JENNY, AND H. A. TCHELEPI, *Multi-scale finite-volume method for elliptic problems in subsurface flow simulation*, J. Comput. Phys., 187 (2003), pp. 47–67.
- [51] S. LEONARDI, *Weighted Miranda-Talenti inequality and applications to equations with discontinuous coefficients*, Comment. Math. Univ. Carolin., 43 (2002), pp. 43–59.
- [52] A. MAUGERI, D. K. PALAGACHEV, AND L. G. SOFTOVA, *Elliptic and Parabolic Equations with Discontinuous Coefficients*, Mathematical Research 109, Wiley-VCH, New York, 2000.
- [53] F. MURAT, *Compacité par compensation*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (4), 5 (1978), pp. 489–507.

- [54] F. MURAT AND L. TARTAR, *H-convergence*, in Topics in the Mathematical Modelling of Composite Materials, Progr. Nonlinear Differential Equations Appl. 31, Birkhäuser Boston, Boston, MA, 1997, pp. 21–43.
- [55] A. A. OBERAI AND P. M. PINSKY, *A multiscale finite element method for the Helmholtz equation*, Comput. Methods Appl. Mech. Engrg., 154 (1998), pp. 281–297.
- [56] H. OH AND I. BABUŠKA, *The p-version of the finite element method for the elliptic boundary value problems with interfaces*, Comput. Methods Appl. Mech. Engrg., 97 (1992), pp. 211–231.
- [57] H. OH AND I. BABUŠKA, *The method of auxiliary mapping for the finite element solutions of elasticity problems containing singularities*, J. Comput. Phys., 121 (1995), pp. 193–212.
- [58] H. OWHADI, *Anomalous slow diffusion from perpetual homogenization*, Ann. Probab., 31 (2003), pp. 1935–1969.
- [59] H. OWHADI, *Averaging versus chaos in turbulent transport?*, Comm. Math. Phys., 247 (2004), pp. 553–599.
- [60] H. OWHADI AND L. ZHANG, *Metric based upscaling*, Commun. Pure Appl. Math., 60 (2007), pp. 675–723.
- [61] G. PAPANICOLAOU AND S. R. S. VARADHAN, *Boundary value problems with rapidly oscillating random coefficients*, in Colloquia Mathematica Societatis János Bolay 27, 1979.
- [62] G. PAPANICOLAOU AND S. R. S. VARADHAN, *Diffusions with random coefficients*, in Statistics and Probability. Essays in Honour of C. R. Rao, N. Ikeda, ed., North-Holland, Amsterdam, 1982, pp. 547–552.
- [63] G. C. PAPANICOLAOU, D. STROOCK, AND S. R. S. VARADHAN, *Martingale approach to some limit theorems*, in Papers from the Duke Turbulence Conference (Duke Univ., Durham, N.C., 1976), Paper No. 6, Duke Univ., Durham, NC, 1977, pp. ii+120 pp. Duke Univ. Math. Ser., Vol. III.
- [64] G. C. PAPANICOLAOU AND S. R. S. VARADHAN, *A limit theorem with strong mixing in Banach space and two applications to stochastic differential equations*, Comm. Pure Appl. Math., 26 (1973), pp. 497–524.
- [65] S. SPAGNOLO, *Sulla convergenza di soluzioni di equazioni paraboliche ed ellittiche.*, Ann. Sc. Norm. Super. Pisa (3) 22 (1968), 571–597; errata, ibid. (3), 22 (1968), p. 673.
- [66] D. W. STROOCK AND S. R. S. VARADHAN, *Diffusion processes and martingales. I*, in Martingales (Report Meeting, Oberwolfach, 1970), Lecture Notes in Math. 190, Springer, Berlin, 1971, pp. 60–66.
- [67] D. W. STROOCK AND S. R. S. VARADHAN, *Diffusion processes and martingales. II*, in Martingales (Report Meeting, Oberwolfach, 1970), Lecture Notes in Math. 190, Springer, Berlin, 1971, pp. 67–75.
- [68] D. W. STROOCK AND S. R. S. VARADHAN, *Martingales. I, II, III*, in Topics in Probability Theory (Sem., Courant Inst. Math. Sci., New York, 1971–1972), Courant Inst. Math. Sci., New York, 1973, pp. 113–161.
- [69] D. W. STROOCK AND S. R. S. SRINIVASA VARADHAN, *Multidimensional Diffusion Processes*, Grundlehren Math. Wis. 233, Springer, Berlin, 1979.
- [70] L. TARTAR, *Compensated compactness and applications to partial differential equations*, in Nonlinear Analysis and Mechanics: Heriot-Watt Symposium, Vol. IV, Res. Notes in Math. 39, Pitman, Boston, MA, 1979, pp. 136–212.
- [71] L. TARTAR, *Homogénéisation et compacité par compensation*, in Séminaire Goulaouic-Schwartz (1978/1979), École Polytech., Palaiseau, 1979, pp. Exp. No. 9, 9.
- [72] S. R. S. VARADHAN, *private communication*, 2005.
- [73] W. L. WAN, T. F. CHAN, AND B. SMITH, *An energy-minimizing interpolation for robust multigrid methods*, SIAM J. Sci. Comput., 21 (1999), pp. 1632–1649.
- [74] X. H. WU, Y. EFENDIEV, AND T. Y. HOU, *Analysis of upscaling absolute permeability*, Discrete Contin. Dyn. Syst. Ser. B, 2 (2002), pp. 185–204.
- [75] X. H. WU, Y. EFENDIEV, AND T. Y. HOU, *Accurate multiscale finite element methods for two-phase flow simulations*, J. Comput. Phys., 220 (2006), pp. 155–174.
- [76] J. XU, *An introduction to multigrid convergence theory*, in Iterative Methods in Scientific Computing (Hong Kong, 1995), Springer, Singapore, 1997, pp. 169–241.
- [77] J. XU AND J. QIN, *Some remarks on a multigrid preconditioner*, SIAM J. Sci. Comput., 15 (1994), pp. 172–184.
- [78] V. V. ZHIKOV, S. M. KOZLOV, O. A. OLEINIK, AND K. T. NGOAN, *Averaging and g-convergence of differential operators*, Russian Math. Surveys, 34 (1979), pp. 69–147.